

# **Three Views of Political Philosophizing and the Idea of State Neutrality**

**CHRISTOS SOFIANOPOULOS**

**Submitted in partial fulfilment of the requirement of the  
Degree of Ph.D.**

**London School of Economics and Political Science**

**November 1995**



UMI Number: U086978

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U086978

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.  
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

THESES

F

7454



623280

## Abstract

This thesis is an attempt to understand the Neutrality Debate, in the light of two basic distinctions: a distinction between political philosophy and politics, and a distinction between three different views of political philosophizing, which I call **foundationalism**, **explanationism** and **interpretivism**.

According to **foundationalism**, the political philosopher starts with an account of what is essentially human and deduces from it moral/political principles, which should govern every human society. According to **explanationism**, the political philosopher seeks to understand a particular society (e.g. the Western state) sub specie aeternitatis, i.e. as an immutable, autonomous, self-sufficient world of ideas. Similarly, **interpretivism** focuses on a particular society (e.g. the Western state), but understands it as an ever-changing world of shared conceptions, understandings and self-perceptions, which are unearthed by the political philosopher. Of the three meta-theories only **foundationalism** and **interpretivism** are normative (regulative), whereas **explanationism** is an intellectual exercise.

What is commonly known as "state neutrality" is actually three different philosophical positions each corresponding to a different one of the three meta-theories. Given that each one of these three state neutralities makes different epistemological assumptions, their combination into a single (meta-theoretically mixed) argument with the intention of making state neutrality more attractive to a wider audience is flawed from a philosophical point of view. Such an argument is "political" in that it seeks to persuade rather than to demonstrate philosophical truth.

# Contents

Preface	5
Acknowledgements	7

Chapter 1 Introduction: The Contextuality of "Neutrality"	8
--	---

## Part I A Meta-Theoretical Distinction

Chapter 2 Three Views of Political Philosophizing	24
2.1 Foundationalism	25
2.2 Explanacionism	31
2.3 Interpretivism	37
2.4 Conflicting Meta-Theories or Different Concerns?	51
2.5 Walzer's "Three Paths in Moral Philosophy"	57

## Part II The Normative Positions

Chapter 3 Foundationalist State Constraint (1): The Millian and Rawlsian Arguments	63
3.1 The Millian State Constraint Principle	66
3.2 Mill's Progressivism and the Constraint Argument	71
3.3 Foundationalism and Rawlsian State Constraint	76
3.4 The Neo-Kantian Essentialist Account	80
3.5 The First Foundationalist Constraint Argument	86
3.6 The Second Foundationalist Constraint Argument	97
3.7 The Three Arguments in <u>A Theory of Justice</u>	100

Chapter 4 Foundationalist State Constraint (2): Conceptions of the Good	108
4.1 The Notion of a Second-Order Volition	110
4.2 Vices as Conceptions of the Good	116
4.3 Culture and Conceptions of the Good	129

Chapter 5 Interpretivist State Constraint	147
5.1 The Direct Interpretivist Constraint Argument	150
5.2 The Constructivist State Constraint Argument	163

## Part III The Neutrality Debate

Chapter 6 State Neutrality	195
6.1 Neutralities	198
6.2 Welfarism: The Other Neutrality	222
6.3 Liberalism and Political Neutrality	231

<b>Chapter 7 The Critique of Neutrality</b>	<b>252</b>
7.1 The Impracticability of Neutrality	253
7.2 Neutrality as a Fraud	256
7.3 The Unacceptability of Neutrality	262
7.4 Flaws in the Philosophical Neutralist Arguments	266
 <b>Part IV The Non-Normative Position</b>	
<b>Chapter 8 Neutrality and the Oakeshottian <u>Civitas</u></b>	<b>278</b>
8.1 Philosophizing About Philosophy and Philosophizing About Politics	278
8.2 Oakeshott as Foundationalist	282
8.3 State Neutrality: The Missing Postulate?	293
8.4 "The Pursuit of Intimations", Politics and Interpretivism	305
<b>Conclusion</b>	<b>313</b>
<b>Bibliography</b>	<b>321</b>

## Preface

The idea of state neutrality is an answer to the question of what should be the moral authority and responsibility of the state. State neutrality holds that the state should be "neutral" towards the different conceptions of the good that its citizens hold. According to many of its advocates, state neutrality is a philosophical position, an implication of philosophical truth.

There are two problems with this view. First, as I argue in Chapter 1, "neutrality" is a semantically overloaded term, a term that is better suited to politics than to philosophy. I, therefore, propose to replace it with "state constraint". Second, there is no single view of philosophical engagement and of the nature of philosophical truth. In Part I, I present three different views of political philosophizing. One which is universalist, essentialist and normative in character (foundationalism), one which is particularist, non-essentialist and normative (interpretivism) and one which is particularist, essentialist, but non-normative (explanationism). Given that each one of these views makes different epistemological assumptions, the philosopher cannot combine them together without becoming a politician, i.e. someone who speaks the language of persuasion rather than that of demonstration. It is quite clear that, despite their claims, neutralists (like their critics) have, indeed, put forward meta-theoretically mixed arguments. In Part II, I present state constraint as two distinct, meta-theoretically

undiluted positions: foundationalist state constraint and interpretivist state constraint. Being non-normative, explanationist state constraint is left out of Part II, as well as of Part III, where I focus on the much politicized neutrality debate, which is, of course, a normative debate. I leave for Part IV the explanationist neutrality of Michael Oakeshott, which, I argue, cannot and should not be implicated in the neutrality debate.

Throughout Parts II and III, I refrain from using the term "liberalism" or "liberal state", so as to avoid unnecessary confusion. The term is introduced only in Part III, where it is identified with a particular institutional arrangement, one which both interpretivist and foundationalist neutralists endorse.

A final point concerning the notes made on the text. Footnotes are reserved for explanatory comments and endnotes for bibliographical references.

## **Acknowledgements**

My first acknowledgement is to Professor Kenneth Minogue who offered me the necessary intellectual stimulation in the early stages of this project. I owe a great debt to my supervisor, Dr Robert Orr, for his observations and encouragement. I am also grateful to Mr John Charvet, who read and commented on a large part of my work. Heartfelt thanks go to Katerina Zacharia for her love and unfailing support. Finally, the debt I owe to my parents goes far beyond this thesis, but without their support and encouragement I would never have completed it.

# CHAPTER 1

## Introduction:

### The Contextuality of "Neutrality"

The terms "neutrality" and "neutral" appear in a number of different contexts and have acquired many different meanings. Hence, before one enters a discussion on the issue of "state neutrality", one is tempted to embark on a semantic exploration. And yet most of those authors who are involved in such a discussion usually assume either that there is a single context-independent sense of "neutrality", which we all share and understand, or that such a sense can be determined through a method of induction.<sup>a</sup>

The main point I want to make in this introductory chapter is that although there are two broad, independent meanings of "neutrality" in the English language, they are not specific enough to be independent of all particular contexts within which "neutrality" appears. In fact both of these meanings are so broad as to require further, context-

---

<sup>a</sup>Only Alan Montefiore, in Neutrality and Impartiality (Cambridge University Press: Cambridge, 1980), pp. 8-22, and Jeremy Waldron, in "Legislation and Moral Neutrality", in R. Goodin (ed.), Liberal Neutrality (Routledge: London, 1989), pp. 63-69, offer semantic investigations of "neutrality" independent of the question of "state neutrality". Authors like Joseph Raz, John Rawls, William Galston and Richard Arneson examine only the semantic ramifications of "state neutrality", although the first two are clearly influenced by Montefiore's account on "neutrality".

dependent, specification. According to the first meaning, neutrality involves "taking neither side in a dispute, disagreement or difference of opinion"<sup>1</sup> and according to the second one, it means lacking those characteristics which are used to place something or someone in one or more (implied or explicitly specified) classes or varieties. Neutrality in the former sense is intentional and causal whereas in the latter sense it is not.

Both of these broad meanings raise questions which are to be answered within the particular context in which the term is applied. With regard to the first meaning, what has to be specified is, first, the identity and nature of the neutral party (i.e. whether it is a person, a state, a group, an institution, an official etc.), second, the identity and nature of the conflicting parties, third, the nature of the conflict (i.e. whether it is a game, a debate, an armed conflict), fourth, the way the conflicting parties are treated by the neutral party and, fifth, the reason why the neutral party is neutral. The last two parameters are clearly dependent on the other three. In other words, they are to be determined within the particular context. As Alan Montefiore notes in his introduction to Neutrality and Impartiality, the neutrality of a referee in a game of football is different from the neutrality of the judge as it is different from the neutrality of a nation not participating in a war between other nations.<sup>2</sup>

The second meaning of "neutrality" raises a rather different question, namely, "What are the characteristics used

to classify non-neutral things (that is, characteristics which the neuter--that which is neutral--lacks)?" or, to put it differently, "In what sense is the neuter incommensurate?". In one particular context, namely in science, "neutrality" is the absence of certain specified characteristics (e.g. having the properties neither of an acid nor of a base) or being at an equilibrium (e.g. lying at the point where the forces of extension and compression meet and offset each other). In biology "neutrality" is the condition of belonging to neither sex. In literature and in everyday language it features as a descriptive term; the neutral object is thought to be lacking any distinctive characteristics, to be colourless or expressionless.

In his discussion of "neutrality" Alan Montefiore overlooks this second meaning and concentrates on "neutrality as taking neither side in a dispute". Although he acknowledges the fact that there are different, context-specific meanings of neutrality, his main objective is to find a definite, context-independent answer to the question of how the neutral party treats the conflicting sides. In his view there are two possible answers: it either assists them (or hinders them) to an equal degree, or it refrains from giving any help (or hindrance) to either of them. Montefiore rejects the latter answer because, as he puts it, "...it is a necessary condition for the concept of neutrality to have any genuine application that it should be open to the potential neutral to choose whether or not to exercise an influence on the situation of conflict."<sup>3</sup>

This point deserves some clarification. Montefiore does not distinguish between two further, secondary meanings of "neutrality". After all, refraining from giving any help (or hindrance) to either side is helping (or hindering) both of them to the same degree. Rather, Montefiore's point is that if one's position makes one's involvement in the conflict impossible, "neutrality" is not the proper word to characterise one's position in relation to the conflict.<sup>b</sup> But surely, this point is conveyed by the term "refraining from giving help or assistance". Refraining from doing something means having the ability to do it, but choosing not to. This notion of self-restraint is, I believe, implicit in the intentional/causal meaning of "neutrality".

Although flawed and inconclusive, Montefiore's argument is rather illuminating. Having defined neutral conduct as "assisting or hindering to an equal degree", he avers that there is a serious problem with this definition: Equal assistance or hindrance may lead to "...strange, counter-intuitive situations in which 'neutrality' is not the word that springs to mind."<sup>4</sup> As an example Montefiore uses the case of two children, each appealing to their father to intervene with his support in a dispute between them. If their father gives equal support to both of them, the older, stronger and more resourceful one is bound to come out on top. For Montefiore this is hardly in accordance with our "intuitive"

---

<sup>b</sup>For example, the fact that Ireland is very far from both Peru and Ecuador and maintains very few or no links with them, does not make it neutral in the recent conflict between them. "Indifference" and "detachment" would be better characterizations of its position.

understanding of "neutrality". But, in truth, it is not in accordance with our understanding of "parenthood". "Neutrality" is not a term which we normally use in connection with family relations and, hence if it is to be brought into this context it has to be given a special meaning; it has to be redefined. In international relations one would hardly ever claim that the neutral should assist the weaker side so as to even up the odds of its being victorious.<sup>4</sup> This clearly means that the meaning of "neutral treatment" is context-specific.

Montefiore overlooks this point and suggests that the problem can be bypassed if we regard neutrality as an option only when there is no clear basis for distinguishing between the initial balances of strength and weaknesses of the parties. He, however, quickly acknowledges that the whole thing would then become "dependent on a...complicated subjective order."<sup>5</sup> He then briefly considers the claim that "there is no coherent application of neutrality at all" only to reject it on the grounds that "there are **various forms** of legal and 'technical neutrality' and certain available conditions of neutrality" which are apparently used all around the world.<sup>6</sup> But instead of drawing from this very fact the

---

<sup>4</sup>It could be argued that this is no longer the case. In the on-going Yugoslav wars it has been argued that the UN should intervene militarily to even up the odds between the stronger side and the weaker side. This means that even within particular contexts one cannot talk of a single meaning of "neutral course of action". As I will claim later on, this is enough reason for us to avoid the use of "neutrality" in the debate about the moral authority and responsibility of the state. If, however, for some reason, "neutrality" cannot be abandoned (which is actually the case since the term features in most of the recent literature), then its contextuality should be emphasized.

conclusion that "neutrality" is context-specific, he takes it as an indication that he should continue his search for a single, independent meaning.

One of the most crucial points made by Montefiore is that "...the adoption of neutrality may very well work in favour of one party to the conflict or to the detriment of another...so long as the neutral does not choose to be neutral in order to favour one side."<sup>7</sup> This point contradicts his view on parental neutrality, i.e. neutrality as evening up the odds of each side's winning the contest, but he brushes this difficulty aside. What he now means to say, is that if the neutral knows that action A will have as a consequence (and not as the intended outcome) action B which will favour one side to the conflict over another, and if he still chooses to follow action A for reasons other than favouring any one side, then he does not lose his neutral status.

As it is usually the case, the basis on which such choices are made is self-interest. Can the neutral party unintentionally (but knowingly) favour one side to the conflict in pursuit of its own interests? Can a particular state, for instance, maintain its trade links with another state involved in a war without losing its neutral status? No straightforward answer is possible. In the second World War Germany tolerated American assistance to Britain and the USSR (before December 1941) so as to avoid drawing the US into the war. Germany recognised the US as neutral because it could not afford to do otherwise. This suggests that neutral, at least in international relations, is the side which is accepted as

such, the side which one or more of the adversaries is not prepared to draw, or not interested in drawing, into the conflict.<sup>d</sup>

A different position is put forward by Joseph Raz who distinguishes between narrow neutrality, according to which the neutral nation is neutral only in those activities and regarding those resources which the adversaries would wish neither to engage in nor to acquire but for the conflict, and comprehensive neutrality, according to which the neutral nation is neutral in all matters including those which are unrelated to the conflict.<sup>8</sup> Implicit here are, first, the view that neutrality does not depend on acceptability and, second, the view that the potential neutral side cannot use its self-interest as an excuse for unintentionally favouring one side over the others.

Now, even if we accept that the neutral party's self-interest does not warrant actions whose (unintended) consequences favour one side to the conflict over another, there may be other grounds on which such actions are warranted. There may be a set of rules or principles which the neutral party subscribes to or abides by and which regulate either the activities of the neutral party alone, or the

---

<sup>d</sup>Even the concept of neutrality found in Public International Law is determined by the interests of the different parties and can, therefore, be understood as "accepted neutrality". In the Encyclopedia of Public International Law (North-Holland Publishing Company: Amsterdam, 1982), s.v. "neutrality, concept and general rules", it is stated that "the laws of neutrality constitute a compromise between the conflicting interests of the belligerents and the neutral State. Their content thus depends on the power relationship between the two."

conflict as a whole. If the latter is the case, then those consequences of the neutral party's actions that have an unequal effect on the adversaries are not in conflict with that party's neutral status or in breach of its neutrality. To use an example, if by penalising one team a football referee brings about its defeat (an unintended consequence since the referee's job is to apply the rules of the game and make sure the opponents subscribe to them), he does not cease to be neutral. Similarly the judge who rules for one litigant does not contradict his neutral position although, unlike the referee, he is not only abiding by rules but making rules. In this case the dispute is concerned with the administration of something much more abstract than a set of game rules. However, both cases have a common, very important feature: the neuter is not merely restrained (i.e. choosing not to enter the fray or to act in ways that would draw him into the conflict), but "constrained" by a set of rules or a moral order whose authority is acknowledged.

It is important to note that in both examples (and in most, if not all, examples that could be mustered) the neutral party can never take part in the conflict in the same way or on the same terms as the adversaries. Unlike the nation which chooses to remain neutral in a conflict between other nations, the judge or the referee can fail to perform adequately, but can never enter the conflict. This means that their neutrality (and the neutrality of the rules or principles they help to administer) is also descriptive (what I have called the second broad meaning of neutrality) in that it connotes the lack of

those characteristics which would make them potential sides to the conflict (e.g. football teams or litigants).

Now, what happens if instead of regulating the conflict (and dictating the role of the neutral), the rules or principles govern only the actions of the neutral party. What if the latter although committed to neutrality, has to abide by principles dictating actions whose (unintended) consequences favour a particular side to the conflict? Does its conduct qualify as "neutral"?

In such a situation the neutral party would not play a specified role within the conflict and would not be considered as internal to it. Consider the following example: Britain declares its neutrality in the Yugoslav wars but adheres to pacifism and regards the maltreatment of civilians and PoWs as an evil. It, therefore, dispatches a peace-keeping force to the area seeking to minimize the conflict and to protect, to the extent that this is possible, the lives of non-combatants. In doing so it occasionally has to obstruct the military operations of one side and castigate its methods. This, however, does not mean that it sides with the opposing sides or that it wishes a particular outcome to the conflict. The question that arises here is whether this kind of limited intervention is justifiable to the warring sides and, if not, whether it contradicts Britain's declared neutrality.<sup>e</sup>

---

<sup>e</sup>One could argue that since the intervention is authorized by the UN, this is a case similar to that of the referee. Nevertheless, the "game" of war is notoriously difficult to regulate and the authority and reputation of the UN as a referee are questionable. For this reason one might argue that Britain's role is dictated by its strategic and economic interests rather than by any principles or rules. The use of

A possible answer is that the reason why Britain or the UN have the right to limited intervention is that the sides to the conflict do themselves appreciate (sometimes openly as their leaders are quoted saying that they deplore the evils of war) of the values of peace and immunity for non-combatants and by acting against them they implicitly acknowledge that they are in the wrong. An alternative answer is that by committing atrocities and expanding the war, the warring factions violate what is essentially human and, therefore, any intervention intended to control such activities is warranted. What is implicit in both answers is the view that neutrality is compatible with limited intervention whose purpose (and result) is not the defeat of a particular side to the conflict.

These or similar arguments are often made in connection with Britain's role in the Yugoslav wars, but are always combined with arguments about Britain's national interests (about the cost of intervention, its relation to economic and strategic interests etc.), about efficiency (whether intervention has actually achieved better conditions for non-combatants and prevented the war from spreading elsewhere), and about national prestige. It, therefore, seems that those values which warrant neutral intervention are hardly the only ones that are taken into consideration and this undoubtedly makes the case for neutrality rather weak. If there are values or principles which are neutral in the "second broad sense" of

---

"neutrality" in this case is, therefore, quite problematic.

the word (i.e. if they lack those characteristics which would make them particularly Muslim, Serb, Croat or British--and certainly a, however qualified, pacifism and the principle of non-combatant immunity **are** such values), it seems that they should not be combined with the vested interests of any one side. In the actual world of politics and international relations this is virtually impossible. Values and principles are always attached to particular vested interests and are used in order to persuade and justify rather than because they are the only true bases for action. It follows that even within this particular context, it is very difficult to put forward a clear definition of neutrality. The tendency will be to regard as neutral the side which is accepted as such by all the sides to the conflict.

"Neutrality" has recently been introduced into moral/political philosophy and is used in connection with the issue of the moral authority and responsibility of the state. The latter is considered to be "neutral" because it is governed by a "neutral" morality (in the second broad sense of "neutral") and because it is not supposed to take action so as to intentionally promote particular "conceptions of the good"<sup>f</sup> or moral outlooks over others (i.e. because it is neutral in the second broad sense of the word).

One might think that since philosophy, unlike politics

---

<sup>f</sup>The term "conception of the good" will be discussed in Chapter 4. For the purposes of this preliminary discussion, it suffices to say that particular values (moral as well as aesthetic) and any considerations which may be thought to have moral implications are included.

and international relations, is not about mustering as many disparate arguments as possible in order to persuade people of the wisdom and desirability of various policies and actions, but seeks to discover the truth and prove what its implications are,<sup>8</sup> it is more likely to have a coherent notion of neutrality. However, the case is that "neutrality" has brought into political philosophy the extra luggage it acquired in various other contexts and especially in politics and international affairs. Since in these contexts neutrality is usually regarded as a means to achieving a desired end, many commentators started to evaluate the "state neutrality" on the basis of what they regard as desired ends. In international relations there are those who think of neutrality as profitable or convenient, those who think of it as a contribution to world peace (which is itself regarded as a good), those who take it to be the opposite of "justice" or of "fighting for what one believes is right" (and, therefore, a bad) and those who regard it as a synonym for "isolationism", as the lack of the power and/or willingness to influence others. Similar weaknesses and advantages were, therefore, attributed to "state neutrality".

Others are not very clear about what exactly distinguishes the "neutral morality" governing the institutional structure of the state from all other moral outlooks and makes it incommensurate with them. They are not, in other words, sure in what sense the state is neutral in the

---

<sup>8</sup>These implications, as we shall see, can be both normative (prescriptive) in character and non-normative.

second broad sense of "neutrality". Some of the advocates of neutrality do themselves add to this confusion by claiming that the neutral state is not governed by a "neutral morality" but by what is not a moral conception at all.<sup>9</sup> In other words, "state neutrality" is equated to "moral neutrality", an idea which was always bound to draw many critics.

To dispel misconceptions of the latter kind one should emphasize that the neutral institutional structure is indeed governed by a conception of the good, although by one which is incommensurate with other conceptions since it is a higher-order<sup>h</sup> moral outlook. After all, as many authors have pointed out, it would be impossible for it to be an a-moral (or morally irrelevant) conception of the good since the very idea of having such a conception is absurd.

Another way of dissipating this kind of misunderstanding is by discarding the term "neutrality". This is an option which should be given serious consideration especially since it would be the only way of dealing with the other problem mentioned above, namely the projection of meanings and connotations of "neutrality" acquired in politics and international affairs on to political philosophy. The "neutrality" which the political philosopher has in mind is not a policy or a course of action whose effectiveness in achieving particular ends can be argued about. It is an implication of philosophical truth. There may be various,

---

<sup>h</sup>One could just as well use the term "lower order morality" to make the same point, namely that the neutral morality is a morality of a different order and that it is incommensurate with the various conflicting conceptions of the good.

disparate bases of evaluation of neutrality qua policy, but the philosopher's notion of neutrality is defined by the theory itself and is not open to the scrutiny of the non-philosopher. If this point is too hard for some to accept, and I believe that this is the case, then "neutrality" should be replaced with a less problematic term.

In Part II of this thesis I will replace "state neutrality" with "state constraint" ("constraint" being, as I have explained, an implication of the first broad sense of "neutrality") in order to avoid any misinterpretations of the neutralist position. In Chapter 6, I re-introduce "neutrality" for the simple reason that it features in virtually all of the arguments that appear in the literature which I examine there. "Neutrality" is then used as a stage-piece without affecting the content of the neutralist position.

I have mentioned that state neutrality is an implication of philosophical truth. The problem, however, is that there is no single view of the nature of philosophical truth. In the next chapter I distinguish between three such meta-theoretical views and go on to say that what is widely referred to as "state neutrality" are actually three different, incommensurate positions, each corresponding to a different view of philosophical truth. This means that the contextuality of neutrality has to be accounted for even within the context of political philosophy.

## Notes

1. Oxford English Dictionary, s.v. "neutral".
2. Alan Montefiore, "Neutrality, Indifference and Detachment", in Alan Montefiore (ed.), Neutrality and Impartiality (Cambridge University Press: Cambridge, 1980), pp.4-17.
3. Ibid, p. 6.
4. Ibid., p. 7.
5. Ibid., p. 7.
6. Ibid., pp. 7-8.
7. Ibid., p. 13.
8. Joseph Raz, The Morality of Freedom (Oxford: Clarendon, 1986), p. 122.
9. Jeremy Waldron, "Legislation and Moral Neutrality", in R. Goodin, A. Reeve (eds.), Liber Neutrality (Routledge: London, 1989), p. 61; David Richards, in Sex, Drugs Death and the Law (Rowman: Totowa, 1982), p. 52; and Michael Oakeshott, On Human Conduct (Clarendon: Oxford, 1990), p. 175 (see section 8.2).

## **Part I: A Meta-Theoretical Distinction**

## CHAPTER 2

### Three Views of Political Philosophizing

As was mentioned in the previous chapter, the main difference between a political and a philosophical argument is that the former seeks to persuade (i.e. to make others accept something as true or false, good or bad, advantageous or harmful, feasible or impossible etc. and perhaps even to cause them to follow particular courses of action) whereas the latter seeks to prove. It follows that a successful political argument is one that is received well by the audience (the particular audience to which it is addressed) whereas the good philosophical argument is that which arrives at the moral and the political truth and/or its implications and is independent of any notion of acceptability. The nature of the truth is defined by what one might call "a meta-theory", "a philosophy of political philosophy" or "a view of political philosophizing".

My aim in the present chapter is to distinguish between three different views of political philosophizing: foundationalism, explanationism and interpretivism. My starting-point is Plato's illustration of the philosopher's engagement in what is known as the Story of the Cave.<sup>1</sup> The imagery of the story alludes to the main issues that underlie all three views of philosophizing and helps to elucidate the

epistemological differences between them.

## 2.1 Foundationalism

In Plato's Story of the Cave, a number of men are chained inside a cave, facing a blank wall, with a fire burning behind them, obstructing the entrance to the cave. These people do not see the light of day and know the world only by the moving shadows it casts on the wall before them. Being in this condition all their lives, they are bound to believe that the shadows are real objects and that they are the whole truth of the world. When one of them is released from his bonds, he goes beyond the burning fire, leaves the cave and discovers the real world illuminated by the sunlight. He later returns and attempts to convince his former companions that what they see on the wall of their cave are but vague reflections of reality.

Without following very closely Plato's explanation of the symbolism in the story, I would say that the cave, with its chained residents and simulacra of the "real world" is a community, a particular socio-political entity with its own practices, institutions, moral and aesthetic values, while the escapee is the philosopher who by leaving the cave abandons the conventional (common) ways of thinking about and doing things. It is this distancing, or withdrawal from "practice" (meaning "convention") that is associated with philosophizing, that is, with the discovery of Ideas (or Forms) of which social meanings, common values and beliefs are incomplete understandings. An Idea, therefore, in this view, is complete

and autonomous, it has an "independent existence".<sup>a</sup> This does not hold for its simulacra in the world of practice and convention. Human nature can have its particular social meanings and can be embedded in custom and convention, but the Idea of human nature is independent of such things and the Platonic philosopher who "sees"<sup>b</sup> the Idea is able to define justice as an independent, universal concept.

The Idea of human nature comprises a number of essential characteristics which can be known philosophically or, to use the imagery of the Story of the Cave, can be discovered outside the cave. Human nature should not be identified with the condition of the members of any particular community, but, as Raymond Plant puts it, should be considered "...against a general metaphysical background which relates to the place of human life and agency in the natural order and to an account of the fundamental purposes and drives of human beings."<sup>2</sup> The Platonic account of human nature comprises three different faculties, namely appetite, spirit and reason.

Returning to the explanation of the symbolism in the Story of the Cave, what has not yet been made clear is the role of the freed, disillusioned prisoner after his return to the cave. Is his acquired knowledge of any use to his former companions? Can it be used in tackling their practical

---

<sup>a</sup>For Plato only the Ideas are without qualification. The objects of our sensory experience are described as just, moral, beautiful, large etc. in comparison with something else.

<sup>b</sup>"Idea" stems from the Greek verb idein which refers to the act of "seeing", but here it means mental rather than physical perception.

problems or is it knowledge of a world that very little resembles their own and can, therefore, have no effect on their lives?

According to one meta-theoretical view, philosophical knowledge, knowledge of human nature or of the essences, enables us to formulate normative (regulative) principles, i.e. evaluative prescriptive rules, which could govern the institutional structure of society or even provide its members with moral guidance in their private lives and/or dictate to them a particular lifestyle. Such principles are founded on knowledge of the essences and are, therefore, derived from, or rather dictated by, them. This point is made by Christopher Berry who states that

[t]he concept of human nature provides a criterion for acting or not acting in the world. This means that the conceptual context within which the facts of human nature are identified is oriented towards practice. Such facts establish a context within which it is possible to identify what is appropriate for humans to do...this factual establishment of what is appropriate is thus also the establishment of a context of normative significance. However, there is here just one context. It is not a question of having the facts of human nature on one side and the values or norms on the other, but rather that the facts and values are best understood as dual elements, like the warp and weft of fabric, that constitute a conceptual whole.<sup>3</sup>

This view is explicitly stated in Platonic philosophy. Unlike Aristotle, Plato contrasts the principles<sup>c</sup> that derive from

---

<sup>c</sup>For Plato the principles that derive from his account of human nature govern both the institutional structure of society and the life of the individual. Given that in different individuals different essentially human faculties predominate, Plato asserts that different individuals are to

the essences with endoxa, i.e. with the principles implicit in the Ancient Greek political culture.<sup>4</sup> The fact that endoxa consists of the imperfect copies of these principles or, to put it differently, the fact that the moral/political principles of any given society are not based on knowledge of the Ideas, means that its members do not know the principles which sustain human nature and are, therefore, in an awkward predicament, the predicament of the chained cave-dwellers. A number of actual problems (perceived as such by the members of the society themselves) may be associated with, or rather caused by this: moral crises, political instability, deep, divisive inequalities, coercion and maltreatment of individuals or groups etc.. Nevertheless, it should be said that the denizens of the cave may not perceive of their situation as a predicament. It is only the philosopher who, having acquired an independent notion of perfection, perceives their situation in this way.

---

occupy different social positions: those whose appetite is dominant should produce, those whose spirit is dominant should defend and those who are predominately rational should rule [Plato, Republic (Penguin: Harmondsworth, 1987), p. 183]. Plato's prescribed arrangements constitute the single, immutable notion of justice, and societies are just only to the extent that they approximate these arrangements.

<sup>4</sup>According to Aristotle endoxa means "communal beliefs", the beliefs that the citizens of the Greek polis actually held on moral/political matters; see Aristotle Nichomachean Ethics (Penguin: Harmondsworth, 1983), p. 40. However, it should be emphasized that the term also refers to principles which are implicit in practice and are thus only tacitly acknowledged. As I will argue later on (see below 5.2) the philosopher should not identify endoxa with expressed opinion (which is often guided by self-interest). It should also be noted that, for Aristotle, endoxa is a particularist, historical specification of foundationalist morality (i.e. of the universally true good). This is not how I use the term here.

It is this normative essentialism that I call foundationalism. In the history of political thought there are quite a few examples of philosophers who found their moral/political prescriptions on what they regard as essentially human. The Stoics, for example, claim that the active material that comprises the whole cosmos, what they called "intelligent pneuma", permeates the whole world and in human adults achieves its purest and fieriest temper as reason. The latter is the crux of the Stoics' account of human nature on the basis of which they prescribe the life of imperturbability.<sup>4</sup> Hobbes and Locke also found their prescriptions on what they saw as human nature. The convention of a "State of Nature" is meant to show that failure to establish a particular social arrangement would compromise what is essentially human. Hume, who challenges ethical rationalism and avers that communal life and, therefore, morality is prescribed by sentiments and more specifically the "sentiment of humanity", is actually putting forward his own account of human nature on the basis of which he justifies a particular institutional arrangement. J.S. Mill proffers an account of human nature which comprises "the diversity of inward forces" and the capacity to be autonomous. In his view, which I will discuss in the next chapter, human nature needs to be protected and this is exactly what the main prescriptive principle of On Liberty purports to do.

What is implicit in all these differing positions is the view that the philosopher withdraws from convention and abandons his parochial interests, his loyalties and sympathies

in order to capture what is essentially and universally human and distinguish it from what is contingent and particular. He then deduces principles of conduct which he presents as dictates of the essences.

For the person who seeks to persuade a wide, diverse audience, a foundationalist argument by itself is not the perfect tool. In Berry's words, its "weakeners seems to be its definitionism", the fact that, in truth, its starting-point is an account of human nature which may not be acceptable to all.<sup>5</sup> In other words, the politician must anticipate and/or accommodate the critics in his audience and this may require combining his foundationalist argument with non-foundationalist ones, or abandoning his foundationalist argument altogether.<sup>6</sup>

Now, foundationalism is not the only essentialist view of political philosophizing. In the next section I will turn to a different essentialist meta-theoretical position, which I will call explanationism.

---

<sup>5</sup>As Berry puts it in Human Nature (Macmillan: London, 1986), p. 93, "it is because any substantive reading of human nature can be challenged that theorists opt for the supposed safer ground offered by formalism... The acceptance of this formalism means regarding 'human nature' as too flimsy a foundation upon which to erect any persuasive or authoritative political programme." It is not quite clear what Berry means by "formalism". The fact that he attributes it to John Rawls (Ibid., p. 90) is not a very strong clue. In my view, it could either mean what I will call "an interpretivist position", or it could mean a mixed foundationalist/interpretivist position, or, again, it could mean some sort of compromise between the differing foundationalist positions (a kind of overlapping area which is acceptable to all).

## 2.2 Explanationism

Unlike foundationalism, which puts forward some account of what is essentially human and deduces from it principles that should govern every human society, the view of political philosophizing I will now examine focuses on a particular society with a particular moral/political culture and seeks to understand it sub specie aeternitatis, that is, to understand it as an autonomous, self-sufficient, immutable world of ideas. The chief exponent of this view is Michael Oakeshott.

In his meta-theory Michael Oakeshott takes the view that philosophical knowledge does not apply to practice and that it cannot be used to regulate or in any way alter the lives of the non-philosophers.<sup>f</sup> In terms of the imagery of the Story of the Cave this means that knowledge acquired outside the cave can neither be used to disillusion the chained cave-dwellers nor to help them get out of any predicament. This is the first fundamental difference between foundationalism and explanationism.

In Oakeshott's view, normativism (which includes what I have called foundationalism) belongs to the world of practice in the sense that it is concerned with the alteration of "what is" into an unrealized idea, a "to be".<sup>6</sup> What is presupposed in this process of purposeful change is, first, that "what is" is a transient, mutable fact and, second, that its outcome is

---

<sup>f</sup>Although Oakeshott is the only self-proclaimed explanationist, he offers a list of thinkers whom he regards as fellow explanationists. The list includes Bodin, Hobbes, Spinoza, Kant, Fichte and Hegel; see On Human Conduct (Clarendon Press: Oxford, 1990), pp. 243-252.

subject to "valuation", or, in other words, that it belongs to the "world of value".<sup>7</sup> The latter is a coherent world of thought which has a certain built-in "objectivity". It is this objectivity that gives practice a normative element by turning "what is to be" into "what ought to be". Philosophical knowledge is not, therefore, contrasted with endoxa or with practice as convention, as in foundationalism,<sup>8</sup> but with practice as action, as purposeful change.

Of course, Oakeshott's moral/political philosophy is more than a negation of normativism. Like the Platonic view, it is explicitly essentialist, but it rejects the universalist character of foundationalism. For Oakeshott, the philosopher's engagement is the pursuit of what is satisfactory (complete) in experience. It stems from a commitment (or propensity) to achieve an absolute coherence, to achieve the concrete whole that is implied in all experiences. In Oakeshott's own words,

philosophical thought [is] the pursuit, for its own sake, of an unlimited unmodified experience, and at the same time [it is] a mood, a turn in mind... [This is] so difficult and dubious an undertaking, leading us so far aside from the ways of ordinary thought, that those may be pardoned who prefer the embraces of abstraction.<sup>9</sup>

This view of philosophy as an unconventional way of thinking

---

<sup>7</sup>This point needs some clarification. Foundationalists do not hold that foundationalist principles are necessarily opposed to the moral/political practices of all societies. They claim that in certain cases there are discrepancies between the two, but in others there are not. When the latter is the case, it is important for the members of society to understand that their practices are in accordance with the dictates of the universal morality that stems from what is essentially human (this, I take it, is Plato's view).

is quite close to the foundationalist view of philosophy as the understanding of something that is already understood in practical terms. The subject of philosophical and conventional understanding is necessarily the same (e.g. justice, the state, the citizen), but the terms of each "mode of understanding" are fundamentally different.

In order to understand Oakeshott's essentialism it is important to examine his theory of understanding. According to Oakeshottian meta-theory, understanding comes about in stages. The first stage, that of "recognition", involves noticing, comparing and distinguishing characteristics. The second stage, that of "identification", involves grouping together characteristics so as to formulate identities. In the third stage, such identities are compared with one another in accordance with fixed, unquestionable criteria which Oakeshott calls "ideal characters".<sup>9</sup> Ideal characters facilitate our lives by providing fixed points of reference, the "embraces of abstraction" mentioned above. The three stages of conditional understanding are, therefore, stages of practical thought. They are part of action or purposeful change.

To use the imagery of the cave, Oakeshott's ideal characters are the moving shadows reflected on the cave's walls and the only common points of reference for its residents. The cave itself is not a prison, a symbol of the human predicament, but a shelter that provides protection from the winds of uncertainty.

Of course, the process of understanding does not stop with the formulation of the ideal characters. Philosophy is

the exploration of their conditionality. In other words, philosophy starts where practice ends and this means that it is both limitless and of no practical use. First, it is limitless because it questions the conditionality of the unquestioning criteria of practice and can go on questioning the new bases of understanding that it establishes. Second, it is non-practical because it is of no use to anyone who wants to achieve an imagined "to be". Philosophy, in this view, can only inspire or provoke new philosophical explorations; it cannot provide the basis for action. This differentiation from the foundationalist view is the crux of Oakeshottian meta-theory.

Of course, Oakeshott does not only explain his view of the philosopher's engagement, of the philosopher's trade, but actually produces philosophical work that is in accordance with that view. To use the non-political idiom, he puts his theory of theory to work. The man of action can therefore deduce a certain methodology leading to Oakeshottian philosophizing. But does not this mean that Oakeshott's view, which is based on the separation of the practical from the philosophical, is flawed? Is Oakeshott refuting his own meta-theory by consciously trying to produce a non-normative form of philosophizing? Is not the production of such a philosophy a purposeful task rather than the result of an inclination, a propensity to philosophize? The answer is No. Oakeshott's meta-theory is not a guide to "how to become a philosopher", but a work of philosophy in itself. Oakeshott does not talk about what he intends to do or what should be done; he does

it. He explores the conditionality of "philosophy". The latter is taken as an ideal character, but its is an idea that non-philosophers can only understand conditionally. "Questioning the conditionality" of such an idea means finding what is essential to it, finding those secondary ideas, its postulates, which are exclusively its own and understanding it in terms of these. The object of the search and study of postulates is to understand the central idea, the ideal character, as a coherent, self-sufficient whole, as a world of ideas.<sup>h</sup>

The same is done in On Human Conduct where the position of the ideal character to be contemplated is the idea of the Western state or "Civitas". Civitas is understood in terms of its own postulates, its essential characteristics, so that all of them together form an autonomous, self-sufficient whole. Oakeshott makes it clear from the start that Civitas is to be understood as an autonomous, self-sufficient entity, as an end in itself.<sup>10</sup> To understand the Western state in this way is to move away from practice, to grasp it as an unadulterated idea which does not correspond to any historical institutional arrangement and can never be realized in the actual world.

The selection of postulates is, of course, one of the philosopher's main concerns. Since the postulates are those features without which the ideal character would not be a world in itself, they should not be in breach of its self-

---

<sup>h</sup>Of course, it should be noted that although this does not involve an idea of human nature, it certainly involves a theory of cognition which is universal. It is this theory that provides Oakeshott with the criterion for the selection of postulates.

completeness. In order to make sure that his postulates meet this condition, Oakeshott contrasts the ideal character in question with other ideal characters. In Experience and Its Modes he contrasts Philosophy with History, Science and Practice, whereas in On Human Conduct he contrasts Civitas (or the Civil Association) with Enterprise Association (or "purposeful association") and with Hegel's Bürgerliche Gesellschaft (or "the economy"). This enables him to circumscribe the area under exploration, avoid the overlapping of ideal characters and select the proper postulates.

What becomes clear in this discussion is that in Oakeshott, as in Plato, the philosopher looks for what is essential or necessary to certain distinct, self-sufficient ideas. But unlike Plato, who attributes these ideas to a different, superior world of formal reason, Oakeshott draws from communal life in Western democratic societies. Concepts like "philosophy" or "the state" are expressed in words we commonly use in our everyday lives. Oakeshott's is, therefore, a different kind of essentialism, one that is based on the assumption that we can know the essences of the abstractions we invent through conventional languages to facilitate our lives, that is, the essences of the conceptual tools we use in practice.<sup>1</sup> And this, of course, makes philosophizing a non-

---

<sup>1</sup>It could credibly be argued that Plato himself engages in explanationist philosophizing, or rather that there is a strong explanationist element in his work. This, however is not the point. After all, as I will argue in Chapter 6, Oakeshott himself has produced foundationalist and interpretivist writings. Rather, the point is that the normative, universalist essentialism which is often regarded as most characteristic of Plato is distinct from the non-normative, particularist essentialism of Oakeshott.

practical, albeit satisfying, enterprise. To use once again the imagery of the Story of the Cave, the philosopher leaves the cave not to discover a different world, but to contemplate the realities of the cave away from it, to the think of the cave as a complete, self-sufficient, immutable world.<sup>j</sup>

### 2.3 Interpretivism

So far I have presented two essentialist views of political philosophizing: one which is universalist and normative in character and one which is particularist and non-normative. I will now focus on a third view which combines two of the characteristics of the other two. Like the first one it is normative and like the second one it is particularist. It is the particularist element of this third view that will be the starting-point of the present discussion. The main idea here is that we cannot and, perhaps, should not philosophize about the moral and the political without adhering to the values, aspirations and ethos of our own particular community. This idea may take the form of either of two arguments. First, foundationalism, the search for human nature and the immutable truths that give it meaning cannot deliver a moral outlook of the richness and density that an actual lived-in morality exhibits and which is necessary for it to function.

---

<sup>j</sup>This is ultimately the reason why Oakeshott's **political philosophy** is non-normative. Although one could argue that he fails to adequately explain why the philosophical mode of thinking has no influence on lived life, one cannot claim that the explanationist view of the Western state *qua* immutable, autonomous world (i.e. Oakeshott's specifically **political philosophy**) can be achieved or can be a realistic objective. (See below 8.1).

Second, foundationalism fails to grasp the moral and the political for these can only be found within a particular community, or to use the Platonic fable, inside the home of the chained cave-dwellers. This means that their cave is not a reflection of the real, inaccessible, immutable world, but is itself the only world there is, a world which is under constant change as are the shapes of the shadows on its walls.<sup>k</sup> The first point is made by Michael Walzer who avers that

We do not have to discover the moral world because we have always lived there... No design procedure has governed its design, and the result no doubt is disorganized and uncertain. It is also very dense: the moral world has a lived-in quality, like a home occupied by a single family over many generations, with unplanned additions here and there, and all the available space filled with memory-laden objects and artifacts. The whole thing, taken as a whole, lends itself less to abstract modelling than to thick description. Moral argument in such a setting is **interpretive** in character, closely resembling the work of a lawyer or judge who struggles to find meaning in a morass of conflicting laws and precedents.<sup>11</sup>

This point against foundationalism gives us a first view of this, quite distinct meta-theoretical position: the philosopher focuses on moral language and practice seeking "social meanings", that is, principles, self-perceptions and general expectations that are latent in them. He, in other words, tries to make sense of the many, sometimes conflicting, elements in a living morality, to find the predominant

---

<sup>k</sup>This means that for the interpretivist the shadows in the Platonic cave must be those of its denizens and not of the pragmata of the outside world.

meanings. His is, therefore, an interpretive task and this view of moral/political philosophy is, quite suitably, called "interpretivism".

The second point against foundationalism is made by Benjamin Barber who states that

While a...[foundationalist] question may take the form: "what are the true necessities [essences] of the world and how do we know them to be so?"...a political question takes the form: "what shall we do when something has to be done that will affect us all and we wish to be reasonable, yet we disagree on means and ends and are without independent grounds by which we might arbitrate our differences?". This formulation makes clear that the real political problem is one of action under conditions of uncertainty, not of truth or justice in the abstract... [T]his viewpoint...eschews metaphysics and circumvents issues of final truth or absolute morals.<sup>12</sup>

In this view foundationalism is a "vicious abstractionism" that threatens to abolish politics, to do away with the moral and political practices through which we understand morally relevant conflict in our society and it, therefore, obscures the truth.<sup>1</sup> The foundationalist offers to solve the problems of uncertainty about the requirements of the shared, lived-in morality in particular cases by abolishing this morality entirely and replacing it with one based on some "absolute truth". Walzer makes this same point alluding to the Platonic Story of the Cave. As he puts it

---

<sup>1</sup>It should be noted that although Barber attacks foundationalism as a meta-theoretical position, he really objects to a particular kind of foundationalism (i.e. neo-Kantianism) which will be the main focus of the next two chapters.

[the foundationalist] must deny the assurances of the commonplace... To what sort of place, then, does he withdraw? He constructs for himself...an ideal commonwealth, inhabited by beings who have none of the characteristics and none of the opinions and the commitments of his former fellow citizens.<sup>13</sup>

According to this argument, leaving the cave is an attempt to abandon the role of the citizen and to view the political through the eyes of an a-political being. Once this view has been achieved, the return to the cave, or rather, the application of the acquired knowledge to life in the cave becomes very problematic. The returning escapee can neither see nor understand politics because political life is life within a particular social practice which cannot be understood in terms other than its own. Foundationalist political philosophy is, therefore, flawed in the sense that it fails to grasp its own subject.

This meta-theoretical argument in favour of normative particularism can also take the form of a political argument addressed to the members of Western democratic societies. According to this argument, foundationalist philosophizing is not only based on false assumptions or leading to erroneous conclusions, but is "dangerous" for the Western political tradition, for our shared way of life. Barber claims that the "assimilation of political judgement to cognition instead of action", is thoroughly undemocratic (i.e. contrary to the Western, Anglo-American political tradition). In truth Barber seeks to discredit foundationalism to his audience by claiming that foundationalist philosophizing is foreign to their tradition and that it can change their preferred way of life.

Walzer directs this same argument against the doctrine of judicial review followed in the United States. He fears that foundationalists may influence Supreme Court judges in their interpretation of the Constitution and claims that when the latter is interpreted in the light of essentialist/foundationalist principles, democracy is undermined.<sup>14</sup>

This critique of foundationalism brings to mind the reaction of the chained inmates in the Platonic Story of the Cave: They take offence at the escapee's rejection of their opinions and criticism of their way of life. What is implicit in their reaction is the meta-theoretical view that the philosopher should not leave the cave but should remain to philosophize about its particulars, to use, so to speak, the already available material.

Having concluded the discussion of the first characteristic of interpretivism, namely its particularism, I will now turn to the second one, namely its non-essentialist character. Interpretivist political philosophy does not seek to discover or determine the essences (seen either as society-independent, external ideas or as the postulates of a particular culture understood as an autonomous, immutable world) since it denies there are any. The question that therefore arises is this: Given that the philosopher does not leave the cave either to find the ideas or to contemplate the conditionality of the particulars of the cave, in what sense is his outlook distinguishable from that of the other cave-dwellers'? In what sense is his engagement philosophical?

Now, it is not my intention to give an independent definition of political philosophizing or to endorse any one of the meta-theoretical positions discussed in this chapter. I more or less regard as political philosophy whatever features in the syllabus of a university course in political philosophy. (And that includes works by exponents of all three positions discussed here.) I assume, however, that philosophical engagement involves some kind of reflective detachment from practice, the examination or formulation of what Walzer calls "general concepts" (like liberty, equality, justice, power or of principles like, say, the Rawlsian difference principle) and that unlike political argument it seeks to prove what is true (either for a specific culture or universally). Does interpretivism, with its particularist non-essentialism, fulfil this minimum requirement?

According to Walzer the interpretivist is concerned with "social meanings" or "shared understandings", that is, with deeply rooted, widely shared moral positions and self-perceptions. These "...are frequently expressed in general concepts--in [society's] historical ideals, its public rhetoric, its foundational texts, its ceremonies and rituals."<sup>15</sup> The interpretivist studies carefully these outstanding features of public life and institutional development in order to identify the general concepts. This, as in foundationalism and explanationism involves examining sets of interrelated ideas which, however, are regarded as social meanings, implicitly or explicitly prescriptive positions dependent on practice, rather than as independent,

transcendental ideas (as in Plato) or as ideas essential to the particular moral/political tradition understood as an autonomous, self-sufficient sphere of ideas (as in Oakeshott). Of course these meanings undergo frequent, sometimes very subtle, changes over time and the interpretivist has to keep track of them and reveal both their content and their normative implications. This involves distancing himself from any loyalties, interests and attachments that he may feel or have and proving that his revelations are the truth about the community's moral and political life. It is, therefore, obvious that interpretivism like the other two views of political philosophizing does fulfil the minimum requirement mentioned above.

Some philosophers have emphasized the dialectical (or dialogical) aspect of the engagement of the interpretivist. Barber claims that in interpretivist philosophy "...theory and practice are reconciled... The criteria yielded by common action are permitted to inform and circumscribe philosophy no less than philosophical criteria are permitted to constrain the understanding of politics and inform political action."<sup>16</sup> If this is simply another way of saying that interpretivism does not look for independent, immutable facts about human nature in the form of general concepts but is concerned with the general concepts implicit in the living practices of an actual society, then it is fully compatible with my taxonomy. However, I believe, that the dialogical aspect of interpretivism is often taken to mean something quite different. For instance, Stanley Kleinberg claims that

political philosophy (meaning what I call interpretivism) refers to two related engagements: First, "...the use of philosophical reflection with the aim of advancing more cogent arguments," and, second, the study of political debate itself in order to learn something about the nature of philosophical argument.<sup>17</sup> What is suggested here is that philosophy and politics merge into a forceful, persuasive kind of argumentation. Through the "study of political debate" the philosopher learns about the interests, attachments and opinions of his audience. He then uses this knowledge to formulate general concepts and applies the principles of deductive reasoning in order to put forward more persuasive arguments. His role, in other words, is to tidy up political arguments and make them more effective. Although this is one way of looking at the political role that political philosophers often assume, it is not related to interpretivism. To reiterate the point, political philosophy does not seek to persuade but to demonstrate, to prove. And this is certainly true of interpretivist philosophy.<sup>18</sup>

This, however, is not the only misconception of interpretivism that is endemic in contemporary political philosophy. Interpretivist philosophy is also confused with communitarianism and traditionalism. Assuming that communitarianism prescribes principles conducive to a closely

---

<sup>17</sup>The case of Oakeshott deserves a mention at this point. In Rationalism in Politics and On Human Conduct he implies that politics and interpretivism (what he calls "the pursuit of intimations") is one and the same thing. As I will argue below, in 8.4, this is not the case. Oakeshott simply confuses interpretivism with a postulate of civitas.

knit community, there is no reason why it cannot take the form of a foundationalist position. Furthermore, there is no reason why an interpretivist should not prescribe non-communitarian principles.<sup>17</sup> As Walzer avers in "The Communitarian Critique of Liberalism", the interpretivist can prescribe individualist principles without contradicting his view of philosophical engagement. In fact he argues that such principles can be shown to be implicit in Western moral tradition, a tradition which, in his view, is at least as individualist as it is communitarian.<sup>18</sup>

Now, the confusion of interpretivism with traditionalism (or conservatism) has more to do with its particularist/relativist character. The interpretivist philosopher prescribes different moral principles to societies with different social meanings. For example, in a society "whose social meanings are intergraded and hierarchical, justice will come to the aid of inequality."<sup>19</sup> Seen in this way, the interpretivist's position appears to be nothing more than a justification of the status quo, an endorsement of the existing distributive arrangement. But of course this is a misreading of interpretivist meta-theory. Existing distributive schemes and institutional arrangements are not necessarily consonant with social meanings. The latter are independent of personal or group interests, claims and expressed opinions and prescribe distributive principles of their own. To illustrate the point I will now turn to the

---

<sup>17</sup>These are often lumped under the labels of "individualism" and "atomism" which sound too extremist, as indeed does "communitarianism".

often cited Walzerian argument for universal health service coverage.

According to Walzer, health care in Western societies is a social good with a particular social meaning. In order to determine the latter and understand what it dictates the interpretivist must examine the language, customs and social history of Western societies. First of all "health care" means the restoration of physical health. It follows that, as a social good, it should be distributed by the state to all those members of society who are unhealthy. Some of the institutional arrangements in the U.S. reflect this very point. There are general vaccination programmes, there are programmes that provide some care to the poor and the elderly and there is public funding of medical research. And yet members of the more affluent social groups have better access to costly, extensive, high-quality medical services while some of the poor are deprived of every form of medical attention. Walzer reads this as a discrepancy between the public conception of medical care and the existing scheme of distribution. In his view the medical profession has always had a "bad conscience" about the link between health care and the market and this is proven by the fact that there have always been doctors who offered their services to those who could not afford them. Even the call "Is there a doctor in the house?" reflects a moral expectation that doctors will hasten to meet emergencies whenever they occur. To consolidate his position Walzer turns to the European history of institutional development and finds that health care was from the very early

years regarded as a need rather than as a luxury to be bought at a price. In his view, until confidence in the possibility of curing disease strengthened, Medieval Christians concentrated in the cure of the soul in which they had more faith. For that purpose there was easy access to churches, the teaching of catechism, confession etc.. When the curing of disease became possible and physical health began to be taken seriously, the licensing of physicians, the establishment of publicly funded medical schools and other such measures succeeded the earlier forms of a publicly provided health care.<sup>20</sup>

So to summarize Walzer's argument, language, custom, history and some aspects of the institutional arrangement give evidence in support of the view that, in the particular societies he investigates, the social meaning of health service requires free, universal coverage. The application of this position in the U.S. would involve broad, expensive reforms which the medical establishment would fiercely oppose. Walzer's interpretation is, therefore, a critique of the present arrangement, of the status quo; not an endorsement of it.

It is quite obvious that Walzer's interpretation ignores all evidence in support of the view that health care in Western societies is assumed to be a marketable good. People in Britain, for instance, buy expensive medical treatment in private clinics although they are entitled to free treatment in public hospitals under the National Health Service. Money buys them fast and effective treatment while publicly provided

health service is slow and inefficient mainly due to lack of state motivation and sufficient state funds. Even if we accept the connection between medical treatment in modern times and absolution in Medieval Europe, it is difficult to ignore the fact that the practice of simony in the Middle Ages ensured the distribution of the medieval equivalent of health care in accordance with wealth.<sup>20</sup>

So it could be said that there are at least two possible interpretations of the social meaning of health service in Western societies. Waltzer uses no external criteria or method showing that one of them is more accurate than the other. In fact he himself avers that no understanding of meaning can be final and definitive and this means that all interpretations are educated guesses of people who cannot know how things will turn out or whether the tendencies they have identified will persist.<sup>21</sup> This clearly is a sceptical position. For the sceptic who subscribes to interpretivist meta-theory, but does not think it is possible to pick a single, correct interpretation on the basis of presented evidence, the relevant question is whether there is some other way to determine the most accurate interpretation. Walzer answers the question by likening social interpretation with literary interpretation. As he puts it

---

<sup>20</sup>The idea of drawing evidence from the past poses a certain problem for the interpretivist. He has to avoid making any suggestion that there are principles which are essential to the moral tradition he investigates or principles derived from universal essences. But how can he look for evidence in the distant past without implying the existence of (either kind of) essences?

[t]he best reading [of a poem] is not different in kind, but in quality, from the other readings: it illuminates the poem in a more powerful and persuasive way. Perhaps the best reading is a new reading, seizing upon some previously misunderstood symbol or trope and re-explaining the entire poem.<sup>22</sup> The case is the same with moral interpretation...

The best interpretation, in this view, must be original, powerful and persuasive or, in one word, impressive. Most importantly, however, it must be a "re-explanation of the entire poem" to a particular audience. As Georgia Warnke avers, the latter point has been the main thesis of literary hermeneutics at least since the time of Schleiermacher: the literary critic must reconcile all the different parts of a text and show how these work together to compose a well-intergraded meaning.<sup>23</sup> It is the idea of a comprehensive unity of meaning that guides the critical understanding and separates the better interpretations from the worse. It is not very clear how Walzer uses this principle in connection with social interpretation. What exactly is the equivalent of the "entire poem" when it comes to social interpretation? Walzer's answer seems to be that it is taking into consideration all information related to the development of the social meaning of a particular good, like health care. This, however, leads us nowhere because the selection of relevant information is the interpretation itself. For instance Walzer does not consider the existence and successful operation of private clinics relevant to the development of the meaning of health service and this is why he reaches the conclusion he does.

A more plausible answer is that the "text" here is simply



the denotation of, say, "health care" and that a successful interpretation consists in exploring its connotation and arriving at a **single** prescriptive, distributive principle like "publicly provided equal distribution" with respect to health service. It follows that an interpretation that arrives at a mixed principle (for example a principle allowing of two parallel schemes of distribution of health care, one public and one private) is suspect. According to Walzer, the view that "...the best interpretation is the sum of all the others, the product of a complicated piece of survey research [is no more credible than the view]...that the best reading of a poem is a meta-reading, summing up the responses of all the actual readers."<sup>24</sup>

This view is not, however, without its problems. If indeed health care has been perceived throughout the centuries both as the satisfaction of a need and as a response to the demand for a luxury, why should we not prescribe a mixed scheme of distribution--pretty much like the one presently in place in Britain, France and a number of other Western European countries? Why would the inclusion of both conflicting perceptions of health care be discreditable? Walzer's answer seems to be that a more inclusive interpretation loses its critical force. It is a fatuous statement rather than an attempt at resolving the conflict between those latent perceptions and expectations which comprise social meanings. In my view this is not the case. Walzer, like virtually all political philosophers, can hardly resist the temptation to exert influence on his audience. As

he unwittingly confesses, he wants his interpretations to be persuasive, captivating. His ambition is not to reveal the truth, but to be influential. And perhaps this is a good enough reason in itself for not accepting his interpretative criteria. His scepticism, or rather, his willingness to appeal to the sceptics in his audience steers him away from interpretivist philosophy towards politics.<sup>9</sup>

## 2.4 Conflicting Meta-Theories or Different Concerns?

Implicit in my presentation of the three different positions on political philosophizing is the view that they are competing meta-theories or philosophies of philosophy. It could be suggested, however, that they are just different concerns, different projects to be undertaken by the political thinker on different occasions. The latter seems to be the view that Michael Oakeshott would endorse in the light of his purist view of the philosophical engagement: philosophy is the exploration of the conditionality of ideas which we commonly formulate and use in order to facilitate our life. Foundationalism and interpretivism either formulate such practical ideas or disclose the hidden meaning of those already in use. But they do not question their conditionality, they do not, in other words, try to understand the particular

---

<sup>9</sup>It should be emphasized that scepticism (the view that we cannot know which interpretation is correct) is not an integral part of interpretivist meta-theory. It should also be noted that there is a great difference between this kind of scepticism (which could be called "interpretivist scepticism") and that of someone who accepts that there are universal essences, but doubts that we can acquire knowledge of them. The latter is discussed in 6.3.

sub specie aeternitatis. Instead, both foundationalism and interpretivism assume that implicit in these ideas is a "to be" which will be achieved only if certain arrangements are in place. This does not mean that foundationalism and interpretivism are somehow flawed, incoherent, undesirable or for some other reason objectionable. It simply means that they belong to the world of **practice** rather than to that of philosophy. The political philosopher may engage in social interpretation or in the quest for the essences that comprise human nature (as Oakeshott himself does in many of his writings) but not as part of his philosophical engagement.

I have already explained in what sense foundationalism and interpretivism are philosophical, non-practical, engagements: Like explanationism they aspire to prove rather than to persuade. But there is yet another sense in which they are non-practical. Although they both prescribe a particular "to be", they are not interested in how the transition from the present state of affairs, from the particular "what is", to the prescribed objective will be achieved. This is a political, non-philosophical issue that is treated, if at all, separately. Of course there is no denying the fact that Oakeshott's conception of the philosophical (or of the non-practical) is different and more exclusive. My point, however, is that there are two senses in which foundationalism and interpretivism, like explanationism, are non-practical. But this does not answer the question of whether the three views are competing meta-theories of different projects for even if they all are philosophical engagements they do not have to be

incompatible with one another.

In Interpretation and Social Criticism Walzer expresses his own view on this issue. He suggests that his interpretivism and what he calls the "path of discovery" (a view of political philosophizing which roughly corresponds to foundationalism) could both yield significant, philosophical findings.<sup>25</sup> Perhaps, he asserts, the principle of utility or the rights of man can be regarded as discoveries, as general concepts discovered outside the Platonic cave. He nevertheless doubts that any moral/political change is due to new morally significant discoveries, to fresh foundationalist positions. In his words, change "has less to do with the discovery...of new principles than with the inclusion under the old principles of previously excluded men and women."<sup>26</sup> This means that foundationalism, unlike interpretivism, can have little effect on our moral life because it ignores the lived-in morality, the thick moral/political language and practices of the particular society. Obviously this is a less severe criticism of foundationalism than the one examined earlier; it is now presented as an approach with slim chances of success rather than as a flawed meta-theoretical position. But, as I have explained, the feasibility or realizability of a philosopher's prescriptions is a political, rather than philosophical, concern. Walzer's statement amounts to saying that a politician can exert more influence by using interpretivist arguments rather than foundationalist ones. But this does not mean that the politician cannot use both kinds of arguments in his attempt to persuade his audience. In fact

he can juggle all kinds of arguments in his effort to be persuasive. But can the philosopher also perform the juggling act without contradicting himself? Walzer fails to answer the question.

In my view there is an important reason why foundationalism, explanationism and interpretivism should be seen as rival meta-theories: All of them answer the same three ontological/epistemological questions in ways that make them mutually exclusive. Only two of the three meta-theories answer any one question in the same way (either in the negative or in the affirmative) and each one of them shares with each of the other two only one answer. The first question is "Are there essences which can be known to us?". Foundationalism and explanationism answer in the affirmative. Interpretivism denies that there are universals that can be known to us and holds that social transition makes the existence of particularist essences (i.e. features which are essential to a moral/political tradition) impossible. It follows that the philosopher can only follow the course or development of the meanings implicit in moral and political practice.

The second question concerns the existence of universals, that is, of morally relevant properties that can be predicated of all the individuals of the class of human beings. According to foundationalism we can acquire knowledge of essences which are also universals. This view is rejected by both the other two views of political philosophizing. For explanationism the essences (human nature included) are ontologically determined by the particular society in which we exist, whereas for

interpretivism no morally relevant essences can be known. Morally relevant can only be a particular society's variable moral meanings.

Finally, the third question is this: Does knowledge of the philosophical findings (whether these are universal essences, postulates of ideal characters or social meanings) prescribe courses of moral and political action? Is, in other words, political philosophy normative? Both foundationalism and interpretivism answer this question in the affirmative. It is relatively easy to see why this is so according to interpretivism: Social meanings cannot but have built-in moral and political principles. For foundationalism, however, the prescription of such principles involves a remarkable conversion of our knowledge of universal essences into principles that could govern the institutional structure of (any) society. To put it differently, for a foundationalist argument to be valid on its own terms, it has to establish a particular relation between its prescriptions (principles and policies) and its essentialist/universalist foundations: It must demonstrate that the former is derived from the latter.

Explanationism's answer to the third question has already been mentioned: by examining the conditionality of the Western state the philosopher creates an autonomous world of ideas (the postulates) which cannot be regarded as a "to be", as a world of experience to be achieved in the future. By setting out a "to be" one initiates a process of purposeful change which presupposes, first, that "what is" is a transient, mutable fact and, second, that its outcome is subject to

"valuation", that it belongs to the world of value. The latter is a coherent world of thought which has a certain built-in objectivity giving practice a normative element by turning "what is to be" into "what ought to be". The philosopher's search for the postulates of ideal characters and the ensuing creation of self-sufficient worlds of ideas are very far removed from this world of purposeful change, from what Oakeshott calls "the ways of ordinary thought".

The three views of political philosophizing are, therefore, mutually exclusive in the sense that each one of them starts with different epistemological assumptions. And yet one can hardly help noticing that virtually all political philosophers (even those, like Walzer and Oakeshott, who are very conscious of meta-theoretical differences) combine all three meta-theories in their work. It, therefore, appears that one must either reject this implication of my taxonomy or assert that most, if not all, philosophical positions ever put forward are inherently flawed. Fortunately this is not a genuine dilemma. As has been mentioned, the combination of disparate meta-theoretical positions is a characteristic of political argument. The fact that political philosophers often juggle arguments of different meta-theoretical backgrounds confirms that they can hardly resist the temptation to assume the mantle of the politician, to use all available arguments in order to persuade their audiences. One must, therefore look for the philosophical in their work and distinguish it from the political.

## 2.5 Walzer's "Three Paths in Moral Philosophy"

To conclude the present chapter I will compare the tripartite distinction made above with one made by Walzer in Interpretation and Social Criticism.<sup>27</sup> Walzer refers to "three distinct paths in moral philosophy": the path of discovery, the path of invention and the path of interpretation. As I have already mentioned, the first one of these roughly corresponds to what I have called foundationalism. In Walzer's words, "...what is involved in [philosophical] discoveries...is something like a dis-incorporation of moral principles, so that we can see them, not for the first time but freshly, stripped of encrusted interests and prejudices. Seen in this way, the principles may well look objective... They exist at some distance from our parochial practices and opinions."<sup>28</sup> Discovery here is equivalent to the exit from the Platonic Cave. What is seen and learned away from it, away from practice, is converted into moral principles which can then substitute or complement the old ones.

Walzer's "path of invention" does not correspond to any one of the three views of political philosophizing presented above. For Walzer moral invention is "the construction of an entirely new moral world..." whose principles and laws "all tend towards the same end."<sup>29</sup> Walzer suggests that there is a strong similarity between this undertaking and Descartes' intellectual project described by its author as an attempt "to reform my own thoughts and to build on a foundation wholly my own."<sup>30</sup> In this view the philosopher strives to invent a "universally inhabitable world, a world for all persons." His

creation derives its force and authority either from the reputation of its creator (an example offered by Descartes himself is that of Lycurgus, the sage who became the lawmaker of ancient Sparta) or from a proper method, a process which yields the right results. In either case, "the point of an invented morality is to provide what God and nature do not provide, a universal corrective for all the different social moralities."<sup>31</sup>

Now if philosophy as invention is universalist in character, as Walzer seems to be suggesting, it must be similar to foundationalism. It is the foundationalist who puts forward a "universal corrective" for all living moralities. It is he who leaves the cave and comes back having acquired a new, different moral outlook. The philosopher-inventor can be thought of as a foundationalist whose "correct method" is a way of distinguishing the essential from the contingent, the universal from the particular.

The problem here is that along with this universalist picture of moral invention Walzer offers a second, particularist one. As he puts it, "Descartes' Spartan analogy suggests a different view, which I think is also Rawls' view, a minimalist version of inventiveness. What Lycurgus creates is not the best city, the city that God would have created, but only the best city for the Spartans, the work, as it were, of a Spartan god."<sup>32</sup> This view of philosophical invention is, as I will argue in Chapter 5, interpretivist. The philosopher interprets the social meanings of a particular society and has them interact with one another in accordance with a process

which bears the characteristics of one of society's major institutions (for example by constructing a hypothetical situation, like the Original Position, which resembles the legislative assembly of a western, democratic society). Walzer does not include this kind of interpretivist invention into his interpretivism (his third path of moral philosophy and the one he endorses). This is the reason why his idea of interpretivist philosophy does not coincide with the third view of political philosophizing presented in this chapter.

Before I close this chapter I should note that I have not addressed the question of whether my list of different views of political philosophizing is exhaustive because it seems to me that within this particular meta-theoretical scheme (the particular set of epistemological questions discussed above) there can be no further alternatives.

I will now go on to argue that there actually are three different positions of state neutrality each corresponding to a different one of the three meta-theoretical views.

#### Notes

1. Plato, Republic (Penguin: Harmondsworth, 1987), pp. 256-264.
2. Raymond Plant, Modern Political Thought (Blackwell: Oxford, 1991), p. 36.
3. Christopher Berry, Human Nature (Macmillan: London, 1986), p. 36.
4. See M. Hadas (ed.) Essential Works of Stoicism (Bantam: New York, 1961).
5. Christopher Berry, supra note 3, p. 90.

6. Michael Oakeshott, Experience and Its Modes (Cambridge University Press: Cambridge, 1989), pp. 256-258.
7. Ibid., p. 258.
8. Ibid., pp. 355-356.
9. Michael Oakeshott, On Human Conduct (Clarendon Press: Oxford, 1990), pp. 3-8.
10. Ibid., p. 110.
11. Michael Walzer, Interpretation and Social Criticism (Harvard University Press: Cambridge, Mass., 1987), p.20.
12. Benjamin Barber, The Conquest of Politics (Princeton University Press: Princeton, 1989), p. 206.
13. Michael Walzer, "Philosophy and Democracy", Political Theory, 9 (1981), pp. 388-389.
14. Ibid., p. 371.
15. Michael Walzer, supra note 11, p. 30.
16. Benjamin Barber, supra note 12, p. 4.
17. Stanley Kleinberg, Politics and Philosophy (Blackwell: Oxford, 1990), p. 11.
18. Michael Walzer, "The Communitarian Critique of Liberalism", Political Theory, 18 (1990), pp. 6-22.
19. Michael Walzer, Spheres of Justice (Blackwell: Oxford, 1988), p. 313.
20. Ibid., pp. 86-91.
21. Ibid., p. 124.
22. Michael Walzer, supra note 11, p. 30.
23. Georgia Warnke Justice and Interpretation (Polity Press: Cambridge, 1992), p. 21.
24. Michael Walzer, supra note 11, p. 30.
25. Ibid., pp. 6-8.
26. Ibid., p. 31.
27. Ibid., pp. 3-32.
28. Ibid., p. 6.
29. Ibid., pp. 10-11.

30. Ibid., p. 11.

31. Ibid., p. 13.

32. Ibid., p. 12.

## **Part II: The Normative Positions**

## CHAPTER 3

### Foundationalist State Constraint (1): The Millian and Rawlsian Arguments

The recent introduction of the notion of neutrality into political philosophy in connection with the issue of the state's moral authority and responsibility has caused great excitement and has divided contemporary political thinkers into neutralists, who share the view that the state should be "neutral" between "conceptions of the good", and anti-neutralists, who reject the notion of state neutrality as flawed, inconceivable or impracticable. The two sides are presently engaged in a fervent debate whose outcome is not as yet decided.

The debate itself is political in the sense that each side musters as many disparate arguments as possible in order to persuade its audience to accept or reject its position on state neutrality. Neutralists, like their critics, combine philosophical arguments of different meta-theoretical backgrounds and often relate them to practical considerations in order to make their position more popular.<sup>a</sup> However, from

---

<sup>a</sup>By "practical considerations" I mean considerations concerning the means of transforming "what is" into a prescribed or desired "to be". These could also be referred to as "political considerations".

a philosophical point of view, foundationalist, interpretivist and explanationist arguments cannot be combined together into a single, compelling case (in favour of or against state neutrality) since each one of the three meta-theories has a different view on the nature of philosophical truth. The philosopher must, therefore, distinguish between three different state neutralities, each corresponding to one of the three views of political philosophizing. This involves extricating the foundationalist, interpretivist and explanationist elements from the mixed arguments offered by the participants in the neutrality debate and turning them into coherent philosophical positions. Of course, the significance of this project should not be overstated: one cannot hope to impose meta-theoretical limitations on a political debate. However, it is important to expose the debate for what it really is, especially because many of those who take part in it profess that their sole concern is to explore the implications of philosophical truth.

As was mentioned in Chapter 1, the political and military analogues of the notion of neutrality are by far the most prominent ones. The introduction of "neutrality" into political philosophy is, therefore, bound to be conducive to the subordination of philosophy to politics. "Neutrality" obscures the meta-theoretical distinction of Chapter 2 by bringing its excess semantic luggage into the issue of the moral authority of the state. It could be said that "neutrality" symbolizes the preoccupation with being persuasive rather than with demonstrating what is true. It is,

therefore, a good idea to replace it before examining the truly philosophical, neutralist arguments.

This move does not involve making any changes to the principles prescribed by neutralist philosophers or to any other feature of their philosophical arguments. The very fact that neutralists can drop the notion of neutrality and still maintain their philosophical positions, strongly suggests that they initially adopted it only in the belief that it would make these positions more attractive to their audience.

The term I propose to use instead of "neutrality" is "state constraint". The reason why "constraint" is better suited to philosophy is that unlike "neutrality" (or at least the most common sense of "neutrality") it does not hinge on the self-interest of the prospective neutral or the interests of the conflicting sides and their willingness to keep him out of the conflict. It is independent of any notion of acceptability or intent. Given that the philosopher prescribes principles which he regards as the dictates of philosophical truth and that he, therefore, is not concerned with their acceptability, he has reason to avoid using a term that brings this parameter into play in connection with one of the prescribed principles.

As was mentioned in Chapter 1, constraint is implicit in a particular sense of neutrality. This is why the term often appears in the literature in connection with state neutrality.<sup>b</sup> Its adoption is not, therefore, likely to

---

<sup>b</sup>For instance Bruce Ackerman uses "neutrality" and "constrained dialogue" interchangeably in his Social Justice in the Liberal State (Yale University Press: New Haven, 1980),

confound all of those who have taken part in the neutrality debate.

My objective in the present chapter is to single out the two most coherent foundationalist state constraint arguments drawing mainly from the work of John Rawls (see 3.3-3.7). My starting-point (see 3.1-3.2) will be the Millian argument of On Liberty which, as I will argue, is the precursor of contemporary foundationalist constraint arguments.

### 3.1 Mill's Constraint Principle

John Stuart Mill's On Liberty is a puzzling and controversial work which purports to defend "one very simple principle". This principle is one of the first--if not the first--explicitly made constraint statements. In Mill's own words, On Liberty is a "philosophic text book [that argues for]...the importance to man and society of a large variety in types of character, and of giving full freedom to human nature to expand itself in innumerable and conflicting directions."<sup>1</sup> This essay makes Mill the forerunner of modern state constraint theories and raises issues that help us distinguish between foundationalist and interpretivist state constraint arguments.

The main question addressed in On Liberty is "What is the

---

pp. 8-12; Andrew Mason regards "neutrality" as an "absolute (or near absolute) constraint" in his "Autonomy Liberalism and State neutrality", The Philosophical Quarterly, 40 (1990), p. 444; Robert Nozick, whose foundationalist position is often regarded as neutralist, talks of "moral constraints on the state" in his Anarchy, State and Utopia (Blackwell: Oxford, 1990), pp. 48-51.

nature and extent of the power which society ought to exercise over the individual?" Mill answers this question with his "simple principle", a principle which "is entitled to govern absolutely the dealings of society in the way of compulsion and control, whether in the form of legal penalties or the moral coercion of public opinion."<sup>2</sup> The content of the principle is given in the following passage:

...the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number is self-protection...[It follows] that the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear because it will be better for him to do so, because it will make him happier, because, in the opinion of others, to do so would be wise or even right.

These are good reasons for remonstrating with him, or reasoning with him, or entreating him, but not for compelling him or visiting him with any evil in case he do otherwise.

To justify that, the conduct from which it is desired to deter him must be calculated to produce evil to someone else... In the part which merely concerns himself, his independence is, of right, absolute. Over himself, over his own mind and body, the individual is sovereign.<sup>3</sup>

Mill's principle of constraint is obviously dependent on a "harm principle". It forbids the use of state, social (collective) or individual action for the suppression of an individual's conception of the good, "physical or moral", when adherence to or pursuit of that conception is not harmful to others.

Mill's principle is, therefore, based on some definition of "harm". In the quoted passage "harm" means "interference

with the others' liberty of action". This interference may not only take the form of physical violence or obstruction but may also become interference with the others' ability to deliberate, to form beliefs, to hold conceptions of the good and to act upon them. This means that Mill's constraint principle is characterized by a certain reciprocity: no one should be persecuted for holding and acting in accordance with a particular conception of the good unless, in so doing, he prevents others from holding and pursuing their own conceptions.

Mill also gives an idea of what harm is not. In his view, doing harm (or "evil") does not consist in disregarding etiquette, considerations of propriety, principles of self-preservation or some conception of prudence. Disregard for such considerations does not warrant any form of coercive action-- undertaken either by individuals or by the state. In Mill's words, the individual is "sovereign over his mind and body".

Nevertheless, there may be reasons to persuade, entreat or remonstrate with someone whose conception of the good is allegedly mistaken, improper, self-destructive or foolish. Mill does not become very specific about the forms that persuasion or entreaty may take. For example, can taxation or subsidization on the basis of held conceptions of the good be part of a process of persuasion? Is state propaganda precluded by the constraint principle? These questions are far more relevant today than they were in Victorian Britain.

To understand the scope and implications of Mill's

principle, it is essential to focus on his justification for its adoption and distinguish the philosophical argument from the political. In my view, the philosophical part of the argument is basically foundationalist and raises all the issues discussed by modern exponents of constraint theories.

In the opening chapter of On Liberty Mill talks about what he perceives as an intolerant society, namely nineteenth-century Britain. His claim is that people show a tendency to impose their own opinions, lifestyles and inclinations on others. He talks of the "despotism of custom" which prescribes singularity and thwarts individuality. As he puts it, "everyone must still dress like other people, but the fashion may change once or twice a year." Change is allowed only on condition that all change together. This imposition of uniformity brings about a world without "...any marked character...[it] maims by compression, like a Chinese lady's foot, every part of human nature which stands out prominently and tends to make the person markedly dissimilar..."<sup>4</sup>

Mill's argument seems to have been sparked off by an enthusiasm for eccentricity. At times it appears that his real concern is the "tyranny" of custom and clothes fashion and this makes his position appear trivial. But it is not so much his motive for writing On Liberty as it is the foundationalist argument itself that gives weight to his position.

Mill's first step is an account of human nature. As he puts it, "[H]uman nature is not a machine to be built after a model, and set to do exactly the work prescribed for it, but a tree, which requires to grow and develop itself on all

sides, according to the tendency of...inward forces which make it a living thing."<sup>5</sup> This diversity of "inward forces" is an essential part of being human. It is this diversity that explains the adoption of different conceptions of the good. It is only natural for people to choose different conceptions of the good for they **are** different. Diversity has, therefore, moral value exactly because it is an aspect of human nature.<sup>6</sup>

In On Liberty, diversity is combined with a certain Kantian element: the individual has the capacity to be autonomous, to be independent of the others' will. This capacity is the second aspect of human nature and, according to Mill, it may never be realized in a dogmatic, intolerant environment where "...human capacities are withered and starved, [where people] become incapable of any strong wishes or native pleasures and are generally without either opinions or feelings of home growth or properly their own."<sup>6</sup> Mill's notion of autonomy is really the capacity to bring out the distinct character of the individual, to release the individual's "inward force", and to acquire preferences, develop conceptions of the good and form original opinions.

The upshot of Mill's argument is that the constraint principle is dictated by morality since it aims at the protection of what is essentially human: diversity and the capacity to act autonomously. This is a higher-order conception of the good in a sense that it is not a detailed

---

<sup>5</sup>It is important to distinguish Mill's "diversity as human nature" from diversity as a social fact. As we shall see in the next chapter, constraint principles can be developed from an interpretation of the political and moral values that are implicit in a diverse society.

picture of an ideal lifestyle, a psychological profile or a list of natural inclinations which precludes particular ideals of the good, but a capacity to seek and fulfil any such ideal. This is why the principle of constraint aims at the establishment of conditions under which constant change and revision of accepted ideals is possible, that is, conditions that guarantee the availability of many, differing ideals, lifestyles, and forms of expression that can be revised and modified to accommodate the diverse nature of human beings.

Thus stated, Mill's argument is quintessentially foundationalist. It is both essentialist/universalist, in that it singles out those characteristics which are essentially human, and normative in that it puts forward a normative principle which draws its moral significance from the essences.

### **3.2 Mill's Progressivism and the Constraint Argument**

The problem, however, with Mill's constraint argument is that it is not as consistent as it is described above. Mill incorporates into the argument two interrelated views which seem to contradict the foundationalist, constraint principle. These two views I shall call "progressivism" and "perfectionism" and I shall claim in this section that although they are more in agreement with Mill's position than it is claimed by the critics, they considerably weaken his foundationalism.

On Liberty is replete with references to "human advancement", "progress", and "the growth of civilization",

ideas which Mill tries to link with the constraint principle. On one occasion he states that "...where not the person's own character but the traditions or customs of other people are the rule of conduct, there is wanting...quite the chief ingredient of individual and social progress."<sup>7</sup> What Mill says here is that when individuals are free to choose, develop and fulfil their ideals of the good, they improve themselves, both individually and collectively. "Progress" here does not only refer to the transition from oppressive uniformity to a situation of state (or social) constraint, where the autonomous choice and pursuit of ideals is possible, but to every single autonomous choice. For Mill every expression of the free will is seen as an improvement, as a choice for the morally better.

According to Mill there are two reasons why autonomous choice is seen as an improvement. First, because it further increases the number of conceptions of the good (given that each individual chooses in accordance with his unique "inward force") and, therefore, brings us closer to the most extensive pluralism possible, and, second, because all the autonomously chosen and developed conceptions of the good are morally superior ones. In Mill's own words, "the only unfailing and permanent source of improvement is liberty...since by it there are as many possible independent centres of improvement as there are individuals."<sup>8</sup>

The former notion of progressivism is in agreement with Mill's foundationalism. Progress as continuously expanding pluralism is in fact required by Mill's account of human

nature. The ideal, according to that account, is for each individual to have his own conception of the good. The problem, however, with Mill's position is that the second notion of progressivism alludes to the existence of non-foundationalist criteria by which to evaluate, or at least characterize, conceptions of the good. Such criteria constitute a higher, nobler, more refined conception of the good. It is this conception that the autonomous individual eventually fulfills in Mill's view through the fulfilment of the ideal of autonomous choice.

But if this conception is not derived from the Millian foundations (and this certainly seems to be the implication of Mill's assertion) how can it be incorporated into the foundationalist argument? There is little doubt that the exalted conception to which autonomous choice eventually leads is (some interpretation of) the endoxa of Victorian Britain which is a contingent moral outlook rather than a system of principles derived from universal essences. The combination of the two distinct notions of progressivism is, therefore, a calculated move on Mill's part, an attempt to make the constraint principle more palatable to his Victorian audience by assuring them that it is fully congruous with the values and principles implicit in their moral practices.

The question that arises at this point is whether Mill's commitment to a contingent, lower-order conception of the good is stronger than his commitment to autonomy and diversity. If this is so, then he should give a different, perhaps more elaborate, account of human nature, abandon the constraint

principle and prescribe a different one, or put forward an interpretivist argument instead.

The incompatibility of the constraint principle with Mill's perfectionist views is the main theme in Susan Mendus's critique of On Liberty. Mendus focuses on what she regards as one particular aspect of Mill's substantive, elaborate conception of the good: his view on marriage expressed in "The Subjection of Women". According to Mendus, "if there is a **correct** ideal of marriage (as he clearly thinks there is), and if failure to embrace that ideal constitutes a prevailing evil (as he clearly thinks it does), then why freedom?"<sup>9</sup>

Strangely enough Mendus picks the wrong example to expose the inconsistency of the constraint position. Mill's view on marriage is autonomy-based and, therefore, in full agreement with the foundationalist, constraint position. He declares that

[w]hat marriage may be in the case of two persons ...between whom there exists that best equality, similarity of powers and capacities with reciprocal superiority in them...is the ideal of marriage; all opinions, customs, and institutions which favour any other notion of it...are relics of primitive barbarism.<sup>10</sup>

The "best equality" that Mill refers to here is equal liberty to act autonomously. This point is made more explicitly in the introduction, where Mill states that "[a]ll women are brought up from the very earliest years in the belief that their ideal of character is the opposite to that of men; not self-will and government, but submission and yielding to the control of

others."<sup>11</sup>

The argument, therefore, is directed against the idea of marriage as an institution based on the subordination of one of the two parties involved; an institution that involves the treatment of women as less than autonomous, independent beings. This, therefore, is not an argument in favour of married life or of the single, "proper" form of marriage, but a corollary of Mill's minimalist foundationalism.

Of course, the progressivist/perfectionist aspect of Mill's argument constitutes its main weakness. If Mill had incorporated it into the essentialist part of the argument, his foundationalism would have been coherent, but he would have to abandon the constraint principle altogether. If, on the other hand, he had omitted any reference to progress and to certain widely shared moral priorities of his time, his argument in favour of state constraint would have been even more fiercely criticised than it actually was.

Mill is not the only philosopher to dilute his foundationalist position with interpretivist elements in order to exert more influence on his contemporaries, but his argument is the precursor of neo-Kantian foundationalist state constraint. His essentialist account, the view that human nature can be compromised if public life is not governed by a higher-order morality derived from the essences, and the constraint principle itself are the main aspects of his position that have been adopted and further elaborated by the neo-Kantians. It is to their foundationalist arguments that I will now turn starting from those of their principal

representative, John Rawls.

### 3.3 Foundationalism and Rawlsian State Constraint

Rawls has not only combined arguments of different meta-theoretical background in order to put together a more persuasive argument, but has claimed that his process of deriving principles governing state action is not based on "any metaphysical doctrine or theory of knowledge" and that it "simply appeals to what anyone can accept."<sup>12</sup> Implicit in this claim is the view that the particular audience he addresses would never find persuasive a foundationalist argument or an argument which (like the argument of On Liberty) incorporates strong foundationalist elements. And yet the argument of his A Theory of Justice has a strong foundationalist component. The constraint principle of A Theory of Justice is actually derived from an account of what is essentially human. This is hardly a new interpretation of Rawls' book. Many critics of Rawls' position, and more notably Michael Sandel, criticize what they have dubbed "the Rawlsian theory of the person" implying that Rawls' prescriptions are founded on an account of what is essentially human. My intention, however, is not to criticize Rawls for contradicting himself or for putting forward a dubious theory of human nature, but to isolate the foundationalist constraint position from the rest of his political argument.

Like Mill, the foundationalist Rawls asserts that there are two different moralities: The one flows from what is essential to being a person while the other is contingent and

may vary in accordance with personal choices, local cultures or fashions. For a foundationalist it is the former that should regulate public life since it is the only one that respects our essentially human characteristics.

Unlike Mill, Rawls takes a very specific view of public life (and the kind of collective action that is predicated by it); he thinks of it as a structure of institutions which are supposed to have the ability to create and convey advantages and disadvantages leading to economic and social success or failure. This "basic structure", as he calls it, is the state.<sup>d</sup> It is, therefore, the state--its structure and function--that ought to be governed by the morality that flows from what is essentially human or from what Rawls calls "justice" or the "right". There is, however, something special about the relation between this morality and the contingent morality (or moralities) that does not stem from the essences: they are, in some sense, compatible. The following general formulation of the Rawlsian constraint principle makes this very point:

The State...does not concern itself with philosophical and religious doctrine [i.e. the contingent morality] but regulates individuals' pursuit of their moral and spiritual interests in accordance with principles...[of justice].<sup>13</sup>

---

<sup>d</sup>It could be argued that Rawls applies a notion of the state that is characteristic of Western, democratic societies and that he, therefore, contradicts his foundationalism. In truth, however, the foundationalist Rawls does not "apply" any already established notion of the state. Rather, he, like Robert Nozick in Anarchy, State and Utopia (Blackwell: Oxford, 1990), derives his particular view of the state from his foundations.

It could be argued that the use of the term "philosophical" in this passage is an indication that Rawls does not understand his argument as a foundationalist project (assuming, of course, that he takes "philosophy" to mean the derivation of normative principles from an account of what is essential to being human). However, "philosophical doctrine" can be taken to mean either the normative principles latent in moral and political practice, or those foundationalist views which are not meant to govern the basic structure and, perhaps, do not have (or were not meant to have) any political implications (say, the philosophy of Schopenhauer and its emphasis on the significance of artistic expression). It follows that "philosophical doctrine" is not on a par with and does not rival the foundationalist, higher-order morality. The application of the latter's principles entails the regulation of the people's pursuit of their conceptions of the good and not the evaluation, promotion or rejection of any such conception. In other words, the principles of justice do not rival the various conceptions of the good. For the foundationalist Rawls the endorsement and application of the morality which derives from the essences does not disprove or threaten the differing, contingent moralities; foundationalist morality refers to the means rather than the ends of individuals' lives. It is this interpretation of the passage quoted above that allows us further to pursue the foundationalist component in Rawls' mixed argument.<sup>6</sup>

---

<sup>6</sup>Alternatively, it could be argued that Rawls does not really mean the distinction between philosophy and his principles of justice and that he makes it only in

Now, the Rawlsian principle of constraint can be stated as a series of premises three of which have a strong meta-theoretical flavour. The series runs as follows: (i) there are moral principles that stem from our knowledge of human nature; (ii) there are moral ideals that are conventional or contingent; (iii) the former are suitable for the regulation of public life (axiomatic); (iv) the latter are compatible with the former and their pursuit can be regulated by them. The last statement is true only because Rawls' account of human nature is such as it is. A different account of human nature (say, that offered by Hegel) would make the essences-derived morality an alternative to many contingent moral outlooks.

Nevertheless, the claim that Rawls' foundationalist morality does not encroach on contingent morality needs to be qualified. Contingent morality is thought to include a wide variety of differing conceptions of the good. The very notion of a conception of the good is, as I will explain later on, defined by the Rawlsian essentialist account. Compatibility is not, therefore, something that we can confirm in terms independent of those specified by the essentialist account itself. The foundationalist morality is a system of higher-order moral values which defines the pursuit of lower-order moral outlooks, or conceptions of the good.

---

anticipation to the critics' reaction.

### 3.4 The Neo-Kantian Essentialist Account

In order to piece together a Rawlsian account of what is essentially human, one needs to focus on the Rawlsian conception of autonomy and more particularly on what Rawls calls the "full" and the "thin" theories of the good. The former refers to the individuals' relation to their ends whereas the latter focuses on certain general wants held by all people and is often regarded as a general theory of motivation.

Unlike the Mill of On liberty, the Rawls of A Theory of Justice is aware of the Kantian element in his work and makes it clear that the Original Position is meant to be an illustration of the Kantian conception of autonomy.<sup>14</sup> According to Kant, autonomy is a property of the will. Autonomous is the will that is not motivated by anything outside itself and which, at the same time, is not abandoned to the lawlessness and disorder of internal impulses and desires. Autonomy is, in this sense, a matter of staying free from external influences and conquering internal disorder by being obedient to a self-made rule. Both the formulation of this rule and its application are identified with the concept of rationality. According to Kant, however, the will cannot be autonomous in the actual, causally determined world. It can only exist timelessly in the noumenal realm. What the foundationalist Rawls attempts to do is take the Kantian conception of autonomy out of the two-world context or, more accurately, to extend it to the actual world and use it to derive a morality (namely, justice) that applies to the basic

structure of society.

For Rawls, the only possible manifestation of autonomous will is rational choice-making and it is the capacity for this that we must take to be the essence of personhood. This point is combined with the view that no particular object of rational choice is essential to being human. In his own words, "...the self is prior to the ends which are affirmed by it; even a dominant end must be chosen among numerous [i.e. contingent] possibilities."<sup>15</sup> A person's ends are, therefore, his solely because he chooses them and not because they are constitutive of him. To put it differently, the self is distinct from and antecedent to any aims, or as Rawls himself puts it, "[t]he essential unity of the self is prior to the choice..."<sup>16</sup>

Rawls explains in great detail the process of choosing aims and formulating rational plans for their pursuit. In his view, a rational plan is "one that would be chosen with **deliberative rationality** from among the class of plans all of which satisfy the **principles of rational choice**... We eventually reach a point though where we must decide which plan we most prefer without further guidance from principle."<sup>17</sup> Upon reaching this point, says Rawls, we can try to "...find a more detailed or illuminating description of the **object of our desires**," hoping that we can then re-apply the principles and reach a final decision.<sup>18</sup>

It is fairly obvious that considerations of "deliberative rationality", "principles of rational choice", "forms of critical reflection" etc. are introduced in order to explain

choice as a process and emphasize the universalist character of choice-making. Although these inventions "do not replace the concept of rationality,"<sup>19</sup> they describe the relation between the essential, individuated self and the various, contingent, available plans.

For Rawls the capacity to choose is a capacity to process information concerning our aims. This processing of information is patterned and the patterns are dictated by "laws of human psychology".<sup>20</sup> Consider, for example, Rawls' so-called Aristotelian principle. According to this, "a person...prefers the [activity] that calls upon the greater number of more subtle and intricate discriminations."<sup>21</sup> According to another principle, that of "inclusiveness", "a plan is preferred to another if its execution would achieve all the desired aims of the other plan and one or more aims in addition."<sup>22</sup> Implicit in these principles is the view that human nature, or even Nature itself, comprises a number of natural tendencies: a natural inclination to master more complex than simpler activities (perhaps as an urge to adapt to one's environment), an inclination to achieve a given end by exercising less than more effort and so forth. It seems, therefore, that convenience, adaptability, economy, effectiveness, continuity and coherence are the objective criteria for critically reflecting on available information and reaching a decision.

Although this process is essentially human, the things chosen, i.e. the specific plans, just like the available information, are not. But how about the aims? What exactly is

their relation to the rational plans and to the self? According to a passage quoted above, the aims are "the objects of our desire". Rawls avers that by "[u]sing the principles of rational choice as guidelines, and formulating our desires in the most lucid form we can, we narrow the scope of purely preferential choice, but we cannot eliminate it altogether."<sup>23</sup> This means that aims are ultimately arbitrary expressions of preference and that choices are made on the basis of sub-rational desires which cannot be evaluated or identified in accordance with objective (rational) criteria. It is hardly surprising that Rawls should make this point. If he were to claim that rational choice is a complete break from the subjectivity and arbitrariness of desires, he would not be able to bridge the gap between the phenomenal and the noumenal worlds of Kantian theory.

It seems, nevertheless, that a person's aims--no matter how arbitrary they may be--are not entirely separate from the plans through which they are pursued. In fact, it would make sense to say that the adoption of a plan is an expression (perhaps the only expression) of our having a desire, a manifestation of our having a set goal. As I have already mentioned, Rawls seems to be aware of this point and at times considers aims and plans to be constitutive of what he calls conceptions of the good.

The Rawlsian emphasis on autonomous (rational) choice means that "personhood" (or the "separateness of being") is an essential feature of human nature. It is this separateness, this moderate solipsism, that leads to the notion of diversity

or "plurality". As Rawls himself puts it "the plurality of distinct persons with separate systems of ends is an essential feature of human societies."<sup>24</sup> In actual societies, autonomous individuals are bound to choose differently, first, because the options and information available to them as well as their physical and mental capabilities vary and, second, because there can be no one "correct" process of rational choice.

Now, we have seen that foundationalist philosophy derives normative moral/political principles from an account of what is essentially human. These principles are meant to regulate the interaction between individuals qua essential characters (i.e. bearers only of the essential characteristics). Given that for Rawls the objects of our rational choices are not essential but contingent or, to put it differently, given that no single adopted lifestyle, moral tradition, ideology or religious doctrine is essential to one's being a person, it follows that foundationalist morality is not derived from any such conception of the good.<sup>f</sup> It also follows that all those features that distinguish one individual from another are of no concern to the foundationalist morality. Social status, natural assets and abilities, intelligence, strength, gender as well as "special features of personal psychology"<sup>25</sup> are not essential to personhood and are, therefore, irrelevant where justice is concerned. This means that justice does not include principles that value specific social or physical characteristics or intellectual abilities.

---

<sup>f</sup>In the Rawlsian view, it is essential that a person holds some conception of the good. This is not saying that a particular such conception is essential to being a person.

One very important aspect of the neo-Kantian essentialist account is the view that all individuals have an **equal** capacity to be autonomous/rational choice-makers. This egalitarian element is not present in all foundationalist positions.<sup>g</sup> It seems, however, that Rawls thinks of it as inherent in his conception of autonomy (along with the notions of personhood, or "the separateness of being", and diversity). Whether this is the case is an issue that need not be discussed here.<sup>h</sup>

Since, the crux of foundationalist meta-theory is the view that the basic structure of society must be governed by principles which protect, comply with or are conducive to the essentially human character, foundationalist neo-Kantianism prescribes principles which are conducive to and respect rational choice-making. As has been mentioned, in the actual

---

<sup>g</sup>For instance, Plato holds that all individuals possess three faculties (appetite, spirit and reason), but that as a fact of nature in different individuals different faculties are dominant. This leads to the division of citizens into three social classes each having different duties and responsibilities. Aquinas holds that all men are equal in being rational, although they have not all actualised their reason to the same extent. This means that some notion of equality **is** inherent in foundationalist positions. So, perhaps, it is more accurate to say that, although some notion of equality is implicit in all foundationalist positions, not all of them share the same notion of equality.

<sup>h</sup>Robert Nozick, whose foundationalist state constraint position is similar to what I will call "Rawls' first foundationalist constraint argument", does not hesitate to present his essentialist account as an array of largely unrelated characteristics. As he puts it in Anarchy, State and Utopia, p. 48, "[t]he...proposals for the important individuating characteristics connected with moral constraints are the following: sentient and self-conscious; rational...; possessing free will; being a moral agent capable of guiding his behaviour by moral principles and capable of engaging in mutual limitation of conduct..."

world of causal relations rationality can be compromised by means of external interference. The constraint principle is, therefore, a principle that protects the individual from coercion and thus realises his essential capacity to be autonomous. Whether such a principle can be deduced from the neo-Kantian essentialist account without the introduction, somewhere along the way, of elements which cannot be traced back to the essences, is an important issue which will be addressed in Chapter 5.

### 3.5 The First Foundationalist Constraint Argument

Although autonomous choice, individual distinctness and metaphysical equality are not the only aspects of Rawls' essentialist account, they are sufficient for a first constraint argument. I shall start with Rawls' formulation of the constraint principle and then connect it with the essentialist account presented above. As J.M. Finnis avers, Rawls' constraint argument consists of two complementary principles.<sup>26</sup> The first is the liberty principle (or first principle of justice) and the second is a harm principle very much like the one put forward by Mill. According to the former,

[e]veryone should have the greatest equal liberty consistent with a similar liberty for others.<sup>27</sup>

And according to the latter,

...justice...requires us to show that modes of conduct interfere with the basic liberties of others or else violate some obligation or natural duty

before they can be restricted.<sup>28</sup>

Starting with the liberty principle, it is the notion of "greatest, equal liberty" that we must try to derive from the Rawlsian essentialist account. The first question that arises is "Why liberty?" The answer to this is rather straightforward. As has been mentioned, autonomy, the most important of human attributes, means the creation of self-made rules and the ability to abide by them without external interference. In the causal world autonomy/rationality requires liberty, i.e. a right to carry out chosen plans and achieve chosen ends without interference from others or from the state. Without liberty the capacity to choose one's ends is compromised and may even be lost. Since foundationalism requires, ex hypothesi, respect for and the protection of the essential character of each person, it must require liberty as a right to choose ends and to act upon them.

It is important to note at this point that Rawls distinguishes between liberty as non-interference, or, to use the Berlinian term, as "negative freedom", and liberty as the ability to achieve, or the chances of achieving, one's ends. In the second sense, the better one's chances, the fewer the obstacles on one's way, the freer one feels. Rawls makes it clear that he does not accept this second sense of liberty. As he put it,

[t]he ability to take advantage of one's rights and opportunities as a result of poverty or ignorance, and a lack of means generally, is sometimes counted among the constraints definitive of liberty. I shall not, however, say this, but rather I shall think of

these things as affecting the worth of liberty, the value to individuals of the rights that the first principle [i.e. the liberty principle] defines.<sup>29</sup>

But, what is the relation between this distinction and Rawls' essentialist account? Is the former congruent with the latter? In answering these questions I should restate three main points in Rawls' essentialist account. First, autonomy involves rational choice-making, that is, the processing of available information and the conception of plans for the achievement of personal goals.

Second, although there are certain general, objective criteria which a rational choice-maker may apply, his aims stem from sub-rational desires. In relation to this point I have claimed that Rawls is inclined to treat the aim and the plan for its achievement as a whole. This means that if actual impediments (like poverty, ignorance and lack of means) do not allow the achievement of a goal (even if this is but a sub-rational desire) then it is dropped in the sense that the agent chooses autonomously--rationally--to abandon it since no rational plan can lead to its achievement. Sub-rational desires are, in this sense, tested before they can be considered as personal goals.<sup>1</sup> Rawls would have us believe that unrealistic goals (desires or even preferences and tastes that are developed over long periods of time) can easily be replaced by more realistic ones. This is noted by Brian Barry who states that for Rawls "people's tastes, aspirations and

---

<sup>1</sup>This point is discussed in more detail in the next chapter.

beliefs are always open to modification."<sup>30</sup>

Third, Rawls states that no particular object of choice is essential to the self and, since justice requires the protection of the essential aspects of personhood (this being an axiom of foundationalist philosophy), there is no reason to compensate for the "inability to take advantage of one's right" to achieve a chosen goal. Only when the inability to pursue the particular goal is due to others' coercive action or to prohibitions and measures dictated by a state-enforced, contingent morality is one's autonomous nature violated and even compromised. It appears, therefore, that the distinction between liberty and the worth of liberty is an implication of Rawls' essentialist account as it has been presented in this section.<sup>j</sup>

Having explained how liberty becomes a requirement of justice, it remains to explain why justice requires the "greatest equal" liberty. We have already seen in what sense individuals are, at the metaphysical level, equal: all distinguishing characteristics which can be used in real life to categorize individuals are contingent. What is essential is their separateness and their capacity to choose autonomously. It follows that if this notion of equality is to be turned into a principle pertaining to real life, it must become "equality of liberty". Given that liberty is essential to the

---

<sup>j</sup>However, Brian Barry suggests that Rawls adopts the replaceability of aims and the "unequal worth of liberty principle" simply because if he did not, the second principle of justice justifying the unequal distribution of wealth, income and power would contradict the liberty principle. ("Chance, Choice and Justice", expanded text of a public lecture given at the LSE on 25 April, 1990, p. 23.)

autonomous, rational individual and that all individuals are essentially equal qua rational choice-makers, "equality of liberty" is the only form of equality required by this kind of foundationalist morality. Had Rawls claimed that individuals do not have the same propensity to choose rationally, that is, that essential characters are classifiable in this way, then justice would require liberty rights, but not "equal" liberty rights.

The equality of liberty prescribed by the liberty principle implies a certain notion of compatibility of all individuals' liberty rights. One cannot exercise one's liberty to the extent that it thwarts the liberty of another. So the extent of personal liberty is determined by this consideration and the foundationalist must, therefore, prescribe the "greatest possible" equal liberty.<sup>31</sup>

Turning now to the harm principle, it should be noticed that it states explicitly what is implied in the liberty principle: modes of conduct which violate others' liberty (their equal rights) ought to be restricted or prohibited. Oddly enough, this principle is in stark contrast with Rawls' official view on the issue of the bigots' status in a just society. The fact that it is not any chosen ends whose choice is prohibited but "modes of conduct" which presumably lead to the achievement of these ends, means that even bigoted conceptions of the good (e.g. racism or chauvinism) are valid objects of autonomous choice. Their pursuit, however, must not violate the liberty principle or endanger the existence of the institutions governed by the foundationalist morality.

Ultimately individuals who choose intolerance find themselves in the situation of those whose liberty is of low value because of a lack of means to achieve their ends. A fascist is, in this sense, somebody with an extremely expensive conception of the good. To achieve his end (e.g. a racially pure society) he may choose to adopt modes of conduct that are likely to be expensive--in terms of money (fines) and time (spent in confinement). On the other hand, he may choose to hold his conception of the good as a belief, a fantasy or a vision.

In a critique of the neo-Kantian position, Brian Barry claims that the reason why neo-Kantians prescribe the state constraint principle is because they include in their essentialist account the view that individuals hold their conceptions of the good in a specific manner, i.e. as beliefs or "personal preferences".<sup>31</sup> He then goes on to criticise this view as flawed, claiming that people do not normally hold their ideals in a detached way but strive to establish them as the overriding morality that governs the basic institutional structure.<sup>32</sup> In one of his earlier essays Barry quotes the following passage in order to illustrate this point:

However bigoted the announcement may sound, the Christian can be satisfied with nothing less than a Christian organization of society--which is not the same thing as a society consisting exclusively of devout Christians.<sup>33</sup>

---

<sup>31</sup>The notion of a "personal preference" is introduced by Ronald Dworkin and is contrasted with the notion of an "external preference", i.e. a preference whose satisfaction involves the limitation of other peoples' liberty.

It is certainly true that certain conceptions of the good are often held as "external preferences" (although I am not sure that this can be said of Christianity as a whole). However, the point is not quite relevant to the issue at hand.<sup>1</sup> Neo-Kantian foundationalism does not assume that any one way of holding conceptions of the good is essential to being human. It prescribes principles which prohibit the treatment of individuals as anything other than equal, autonomous, rational beings but its notion of rationality does not presuppose that individuals hold their conceptions of the good as personal preferences. Rather, the treatment of all conceptions of the good as personal preferences is a dictate of the neo-Kantian notion of equality. Conceptions (like nationalism or religious fundamentalism) whose fulfilment involves the suppression of the autonomous choices of others are regarded as personal preferences and are not suppressed unless their adherents violate (the foundationally derived) criminal law.

But is this interpretation in accordance with what Rawls actually says in A Theory of Justice? Looking at the harm principle, we see that state interference is required not only when modes of conduct violate the liberty of others, but also when they are in violation of "some obligation or natural duty". According to Rawls, obligations and natural duties are "principles for individuals". As he puts it, "[a] complete theory of right [i.e. of foundationalist morality] includes a number of moral principles as well [as principles for the

---

<sup>1</sup>In section 3.7 below, I will explain why certain critics have reached this particular interpretation of the neo-Kantian position.

basic institutional structure]."<sup>34</sup>

Apparently this alters everything said so far about Rawls' position. If foundationalist morality includes a number of moral principles by which one ought to live one's life, then justice becomes a rival morality that conflicts with conceptions of the good, traditional moral principles, religious ideals etc.. The bigot is no longer free to hold ideals on condition that he does not harm others, but has to give priority to justice in his own private life and refrain from choosing certain aims. Justice, to put it differently, becomes an aim to be pursued personally as well as collectively; it becomes a constant preoccupation. The notion of liberty that is consonant with such a view, is no longer liberty as non-interference, but Berlinian "positive freedom".

This, however, is not what Rawls has in mind when he talks about principles for individuals. He states that "...obligations presuppose principles for social forms. And some natural duties also presuppose such principles."<sup>35</sup> They are in a sense corollaries of the principles of justice. One has an obligation to support a just institutional arrangement when "...one has voluntarily accepted the benefits of this arrangement."<sup>36</sup> This notion of responsibility seems to be consonant with Rawls' view of rationality and, therefore, it finds a place in Rawls' moral prescriptions.

But how about those rational duties which do not presuppose principles for social forms? Rawls makes clear that most natural duties apply to us "without regard to our voluntary acts." Thus "we have a natural duty not to be cruel,

and a duty to help another, whether or not we have committed ourselves to these actions."<sup>37</sup> This view is problematic in that it assumes benevolence and altruism to be elements of the essentialist account. If this assumption is adopted then there simply is no need to have a harm principle as such because there is no middle ground left between doing harm and not doing good. This means that the foundationalist morality becomes a rival of many moral outlooks and loses its status as a higher-order morality to be predicated of the basic social structure.

This, however, is not the only problem with Rawls' position. If talk of natural duties seem to be incongruous with the state constraint principle, Rawls' remarks on the agent's psychological motivation totally contradicts it. As he puts it in the closing chapter of A Theory of Justice,

...[w]e cannot express our nature by following a plan that views the sense of justice as but one desire to be weighted against others. For this sentiment reveals what the person is, and to compensate it is not to achieve for the self free reign but to give way to the contingencies and accidents of the world.<sup>38</sup>

What Rawls appears to be establishing here is the priority of the foundationalist morality over the various contingent moralities. But he does so in a self-contradictory way. What he now avers is not only that the foundationalist morality is predicated of the basic structure of society, but that the desire to affirm and act in accordance with this morality is essential to being human. Implicit in this point is that

individuals express their essential human character by not adopting conceptions of the good (say, neo-Nazism or Stalinism) which reject the principles of justice. It follows that those who do affirm such conceptions are but lesser human beings, perhaps equivalent to the Aristotelian "natural slaves". This, I believe, is an implication of the application by Rawls of the Kantian categorical imperative. The autonomous individual chooses his conception of the good under the conviction that the process of rational choice-making (and not the chosen conception) sets rules (i.e. the principles of justice) for others to follow. This conviction becomes an essential characteristic of being human. The problem with this move, however, is that it is incompatible both with the already presented version of the harm principle and with a second version also featuring in A Theory of Justice according to which,

...[w]hile an intolerant sect does not itself have title to complain of intolerance, its freedom should be restricted only when the tolerant sincerely and with reason believe that their own security and that of the institution of liberty are in danger.<sup>39</sup>

Rawls does not explain (at least not in strictly foundationalist terms) why the bigots, who are presumably non-autonomous inferiors, should hold the same liberties the rest of the members of society hold. If the simple fact that they hold intolerant conceptions of the good makes them lesser beings, why should they not be repressed and reformed even when they refrain from acting harmfully or when their

collective power is not enough to endanger the existence of just institutions? Rawls can hold on to his harm principles only if he abandons the particular application of the categorical imperative which divides people into humans and less than humans. What warrants the restriction of the bigots' liberty is, first, the violation on their part of the harm principle and, second, the meta-theoretical axiom that the state must be governed by the foundationalist morality. This means that Rawls needs to adopt a wider conception of autonomy, one that allows for autonomous individuals to adopt subversive conceptions of the good.

The point just made does not, of course, disguise the second main weakness of foundationalist political philosophy (the first being the problem of specification of principles without the introduction of interpretivist elements), namely the fact that there will always be cases of individuals whose bearing of the essential characteristics is doubtful. Bigots, drug addicts, foetuses and those with mental illnesses are among the group of individuals whose status, i.e. their essential humanness, is open to question in many modern societies. Ultimately, the foundationalist can be thought of as someone who determines who will be counted in and who will be left out of the institutional structure governed by the principles of foundationalist morality. I suspect that this is a role which the neo-Kantian foundationalist does not feel very comfortable playing. The least he can, therefore, do is abandon the particular application of the Kantian categorical imperative and accept a more inclusive conception of autonomy.

### 3.6 The Second Foundationalist Constraint Argument

As I have already mentioned, Rawls further elaborates his conception of rationality claiming that it involves an essential concern for certain general desires or "desires for primary goods."<sup>40</sup> These goods are actually conditions required for the pursuit of any particular goal. Rawls maintains that primary goods are certain rights and liberties, opportunities and powers, income and wealth and--something assumed to be made possible in the realization of the other--self-respect.

This elaboration of Rawls' essentialist account provides the basis for a second constraint argument. In the argument presented earlier, liberty is seen as a requirement of autonomy, a condition necessary for autonomous choice-making, itself an essential aspect of personhood. The second argument is based on the assumption that liberty, along with power, wealth and self-respect is an object of our essential wants. This further assumption yields both the constraint principle and the difference principle (which refers to the distribution of opportunities and resources) and broadens the range of the state's authority. Since state constraint is now linked to the distribution of certain goods and assumes a different, more elaborate meaning. Ultimately, the difference between the two arguments translates into the difference between a minimalist-state, right-wing political agenda (pretty much like the one defended by the foundationalist Nozick in Anarchy, State and Utopia) and the left of centre politics that is usually

attributed to Rawls.

As we have seen, Rawls maintains that the initial distribution of all physical and social characteristics is arbitrary since no-one chooses or has any control over who his biological parents are going to be or which social environment he will brought up in. This means that social position, physical attributes and intellectual capacities are 'undeserved' and should not determine the level of wealth, income, power and self-respect that individuals attain. Justice, therefore, requires that corrections are made to offset inequalities which cannot be traced to individuals' autonomous choices. Given that justice treats all individuals as essential characters, in the sense that it seeks to protect their essential characteristics, and that all individuals have the same essential want for primary goods (implicit in their essential capacity to choose rational plans of life), justice must counteract actual, undeserved inequalities. Whereas in the first constraint argument equality means equal liberty rights for all individuals, in this argument equality is taken to mean that all individuals should enjoy the same level of primary goods unless their autonomous choices--made after undeserved inequalities are corrected--lead to an unequal distribution. This surely is a radical interpretation of metaphysical equality and one that is impossible to implement. Rawls, therefore, settles for a principle which accepts undeserved inequalities but requires that "social and economic inequalities are to be arranged so that they are to the greatest benefit of the least advantaged..." individuals.<sup>41</sup>

This is Rawls' second principle of justice.

How Rawls gets from equality of primary goods to the second principle of justice, is outside the scope of this paper.<sup>m</sup> What is of interest here is that this same argument also yields the constraint principle (i.e. the equal liberty and harm principles).

For Rawls it is only income and power that cannot be redistributed equally by the institutional structure of society. Liberty can indeed be allotted equally since it is distinguishable from the "worth of liberty", that is, from the possession of the means and knowledge needed to achieve an autonomously chosen end or carry out a plan. It is implied here that although the satisfaction of our desire for wealth and power depends on our non-essential attributes (like physical strength, appearance and social position) and, therefore, requires corrective or compensatory, action by the state, the satisfaction of liberty does not. This is why this line of argument yields both principles of Rawlsian justice.

This argument also proffers a more complete explanation of the harm principle. It is not only the violation of others' rights and of the rules, laws and regulations of the basic institutional structure of society that constitutes harm, but the breach of obligations one undertakes by being a recipient of the State's compensatory policies. The added emphasis on

---

<sup>m</sup>To reach the particular form of inequality prescribed by his second principle Rawls makes further essentialist assumptions. These are: risk aversion, preference for more cooperative product (i.e. preference for larger unequal shares rather than smaller equal ones), motivation (leading to higher levels of cooperative output) through knowledge that unequal distribution of primary goods is in accordance with justice.

obligations may warrant more state interference in matters of morality. Holding an ideal of the good that exhorts the substitution of the present institutional arrangement with another while at the same time being the recipient of some form of compensation for 'undeserved' inequalities, may be regarded as a breach of a tacitly undertaken obligation even if there is no actual legal violation.

It is important to bear in mind that the distinction between the two arguments is not made by Rawls himself. However, since both arguments are implicit in his philosophy and lead to recognizably different social agendas, they deserve to be distinguished from each other. In the first argument liberty is the only necessary and sufficient condition for rational choice-making, whereas in the second it is one of a number of goods satisfying certain essential wants implicit in the capacity to form rational plans of life.

### **3.7 The Three Arguments in A Theory of Justice**

In the preceding sections I have attempted a foundationalist reading of Rawls's constraint position. The result has been the identification of two parallel arguments in favour of a principle of state constraint. As I explained in Chapter 1 and in the introduction of the present chapter, a state constraint principle imposes an absolute constraint on the moral authority and responsibility of the state. In this section I turn to three arguments which are not derived from an account of Rawlsian essentialism, but are offered by Rawls himself as the actual arguments in defence of his principle. The first

two are connected with the arguments presented above, whereas the third, unlike the other two, is interpretivist in that it draws from our beliefs and experiences as members of democratic, pluralist societies.

Before I present these constraint arguments I should make two important points. First, Rawls presents his essentialist view in the form of a hypothetical situation, the Original Position, in which essential characters (bearing what appear to be the essential characteristics) assemble in order to decide on what the principles of justice are. The underlying idea presumably is that since autonomous, rational choice is the primary essential feature of personhood, justice must be regarded as the object of autonomous, rational choice; not of the choice of 'real' people--with given social positions, physical attributes, intellectual abilities etc.-- but the choice of essential characters. The result is an apparatus that yields principles of justice, an invention that is both impressive and convenient. Rawls manages to enliven his argument and make it more persuasive to his Western audiences by bringing in certain conceptions which are widely shared in Western, democratic societies without dropping his foundationalist outlook.

Second, Rawls, unlike Mill, understands his state constraint position to be the exact opposite of perfectionism. According to him, perfectionism is the view that the state has the moral responsibility to foster the good, the well-being, the flourishing and excellence of all its citizens and to discourage them from actions and dispositions which are not

consonant with its ideals of excellence and well-being. To reiterate the central point of this foundationalist constraint position, perfectionist ideals are contingent and are, therefore, unsuitable for the regulation of the basic institutional structure. On the other hand, justice is a higher-order moral outlook that does not rival the contingent perfectionist ideals adopted by autonomous choice-makers, but proscribes the treatment of individuals as less than autonomous, rational choosers of conceptions of the good.

Let us now turn to Rawls' official constraint arguments formulated in terms of the Original Position. According to the first, the essential characters cannot choose perfectionist standards as a "workable basis" for shifting principles governing state action, because

[to] acknowledge any such standard would be, in effect, to accept the principle that might lead to a lesser religious or other liberty... They [persons in the Original Position] cannot risk their freedom by authorizing a standard of value to define what is to be maximized by a teleological principle of justice.<sup>42</sup>

This means that the acceptance of non-foundationalist, contingent principles limits liberty and since liberty is essential to autonomous rational choice (first argument) or the object of an essential want common to all (second argument), non-foundationalist principles are not consonant with human nature and are, therefore, rejected. To put it briefly, perfectionist ideals cannot be predicated of the basic structure of society because they restrict liberty.

The second of Rawls' actual state constraint arguments goes as follows: the people in the Original Position cannot accept non-foundationalist principles as principles of justice because

[t]hey cannot take chances with their liberty by permitting the dominant religious or moral doctrine to persecute or to suppress others if it wishes...[T]o gamble in this way would show that one did not take one's religious or moral convictions seriously.<sup>43</sup>

Rawls seems to be making a further essentialist assumption here. He seems to be saying that essential characters (the characters in the Original Position) exhibit an intractable individualism in that they choose on the conviction that the conceptions of the good they will adopt in the actual world will be valuable simply because they will be their own. An implication of this assumption is that essential characters hold their conceptions of the good as personal preferences and do not care to convey them to others. This is exactly the point made by Brian Barry (see 3.5): For Rawls an essential aspect of being human is holding one's conception of the good as an opinion or preference without transmitting it to others or putting it forth as true or good or valid and, therefore, as suitable for governing state action.

Like Barry, Finnis, claims that "when certain religious and moral convictions precisely were taken with great seriousness, rational people were indeed willing to admit the perfectionist principles and thereby 'gamble' that the right (from their own point of view) religious and moral beliefs

would be enforced."<sup>44</sup> Moral convictions, in Finnis' view, are not held in the way Rawlsian essential characters hold them and, therefore, Rawls' essentialist account, the foundation of the Rawlsian position, is seriously flawed.

This is a misinterpretation of the Rawlsian argument, but is, in some sense, justifiable because although the Original Position is supposed simply to be a colourful way of presenting the foundationalist position, the second official constraint argument indicates that it is not without serious problems. The participants in the Original Position know neither their identity nor any conceptions of the good which they could choose from. In this sense, their autonomy is different from the autonomy Rawls regards as an essentially human attribute (see 3.4) and which his essential characters assume upon their arrival in the real world. In other words, the essentialist account does not coincide with the set of characteristics of the participants of the Original Position. Those who overlook this point, are bound to find that Rawls incorporates atomism (or mutual disinterestedness) into his foundationalist argument.

The third and last of Rawls' actual arguments is not foundationalist in nature. It is based on a view which is widely shared in Western democracies and asserts that "criteria of excellence are imprecise as political principles and [that] their application to public questions is bound to be unsettled and idiosyncratic."<sup>45</sup> Criteria of excellence are, therefore, regarded as conceptions of the good which have not been tested as political principles and are contrasted with

the western, liberal, political tradition which, however, must be but another contingent, perfectionist ideal to the foundationalist philosopher.

Throughout this section I have referred to "conceptions of the good" as the objects of individuals' choices. In order to examine certain implications of the two foundationalist constraint arguments, namely the choice of self-destructive lifestyles and the significance of cultural membership to the individuals' autonomous lives, it is important that I examine more closely the notion of a conception of the good.

#### Notes

1. J. S. Mill, On Liberty (Penguin: Harmondsworth, 1979), p. 68.
2. Ibid., p. 68.
3. Ibid., p. 69.
4. Ibid., p. 135.
5. Ibid., p. 123.
6. Ibid., p. 125.
7. Ibid., p. 86.
8. Ibid., p. 104.
9. Susan Mendus, Toleration and the Limits of Liberalism (MacMillan: London, 1989), p. 64.
10. J. S. Mill, The Subjection of Women (Virago: London, 1983), p. 177.
11. Ibid., p. 27.
12. John Rawls, A Theory of Justice (Oxford University Press: Oxford, 1990), p. 213.
13. John Rawls, supra note 12, p. 212.
14. Ibid., p. 253.

15. Ibid., p. 560.
16. Ibid., p. 563.
17. Ibid., p. 551. (Emphasis mine).
18. Ibid., p. 551. (Emphasis mine).
19. Ibid., p. 552.
20. Ibid., pp. 136-137.
21. Ibid., p. 412.
22. Ibid., p. 412.
23. Ibid., p. 552.
24. Ibid., p. 28.
25. Ibid., p. 137.
26. John Finnis, "Legal Enforcement of 'Duties to Oneself': Kant Versus Neo-Kantians", Columbia Law Review, 87 (1987), p. 434.
27. John Rawls, supra note 12, pp. 327-328.
28. Ibid., p. 331.
29. Ibid., p. 204.
30. Brian Barry, "Chance, Choice and Justice", expanded text of a public lecture given at the LSE on 25 April, 1990, pp. 22-23.
31. John Rawls, supra note 12, p. 391.
32. Brian Barry, "How Not to Defend Liberal Institutions", in R. Douglass et. al. (eds.), Liberalism and the Good (Routledge: New York, 1990), pp. 39-42.
33. Brian Barry, A Political Argument (Harvester & Wheatsheaf: London, 1990), p. 77.
34. John Rawls, supra note 12, p. 108.
35. Ibid., p. 110.
36. Ibid., p. 112.
37. Ibid., p. 114.
38. Ibid., p. 575.
39. Ibid., p. 219.

40. Ibid., p. 93.

41. Ibid., p. 283.

42. Ibid., pp. 327-328.

43. Ibid., p. 207.

44. John Finnis, supra note 26, p. 454.

45. John Rawls, supra note 12, p. 330.

## CHAPTER 4

### Foundationalist State Constraint (2):

#### Conceptions of the Good

Conceptions of the good are the objects of individuals' choices. A conception of the good is a specific activity or sum of activities that conforms to a certain plan and is derived from a set of priorities, or, to use a more commonly used term, a system of values. A lifestyle, a general outlook, an occupation, a favourite pastime are all regarded by the neo-Kantian foundationalist as conceptions of the good. Straightforward though it may sound, this definition raises a number of questions which the exponent of foundationalist state constraint has to address.

As it has already been mentioned, Rawls regards the choice of conceptions of the good as a process. It is this process that shows what a conception of the good is: First, the individual describes "the object of [his] desires"<sup>1</sup> by applying certain "forms of critical reflection"<sup>2</sup> and, then, processes the information available to him in order to devise a plan that will help him attain the object of his desires. In the latter stage, he follows certain "principles of rational choice",<sup>3</sup> which simply means that he (as a rational choice-maker) is disposed to take the most effective, economical and

convenient way possible under any given circumstances. But, although there may be certain publicly agreed on criteria for the evaluation of the actual execution of the plan, there is no one correct or definitive process of choice-making. By adopting or devising a rational plan, one sets a self-made rule, a pattern of purposeful action which although it may be identifiable, it need not always be fully accessible or intelligible to others.

So choosing a conception of the good involves adopting or devising and following a rational plan of action. But what about the earlier phase that starts off the process? What exactly initiates the process of choice-making? As has been mentioned, for Rawls the starting-point is a sub-rational desire, a want or preference. Having desires, however, is not enough for choosing rationally. One has to articulate them, to formulate one's desires "in the most lucid way [one] can."<sup>4</sup> This involves making comparisons, establishing sets of priorities, and making coherent what may be intermittent, impulsive or whimsical. It is in this lucid way that we can understand and communicate our desires and seek their satisfaction.

Two crucial questions arise at this point: First, how can one be regarded as an autonomous agent when one's choice of plan is, to a large extent, determined by one's sub-rational desires? If being autonomous is overcoming internal disorder, how can we deem autonomous the person who simply responds to his desires? So far I have claimed that it is through articulation-planning-execution that one exhibits one's

autonomous-rational nature.<sup>4</sup> In the present section I will attempt to further clarify this point.

The second question points to a certain ambiguity in the desire-based view of rational choice-making: In everyday life people make choices in accordance with principles or ideals which are independent of their personal desires and preferences. Such ideals "...discriminate among want-satisfactions, assigning a greater value to some than to others and perhaps assigning to some a zero or perhaps a negative value."<sup>5</sup> Are we to conclude that ideals are not adopted in order to satisfy particular preferences and that, therefore, actions performed in accordance with them are non-autonomous?

#### 4.1 The Notion of a Second-Order Volition

To illustrate the answer to both these questions, I shall make use of a much quoted passage by Ronald Dworkin where he states that

...the scholar who values a life of contemplation has a conception of the good, but so does the television-watching, beer-drinking citizen who is fond of saying "That's life", though of course he has thought less about the issue and is less able to describe or defend his conception.<sup>6</sup>

In this example the bibulous TV-viewer has chosen a conception

---

<sup>4</sup>This is the view that Robert Nozick takes in his rendering of the first foundationalist state constraint argument. He avers that we are rational in that we are "capable of using abstract concepts, not tied to responses to immediate stimuli" (Anarchy State and Utopia, p. 48).

of the good that satisfies particular wants or desires, whereas the scholar is said to "value" the life of contemplation, i.e. to have adopted it because it is valuable in itself. The scholar is, in other words, thought to have access to certain evaluative criteria which are independent of any wants or desires he may have. Dworkin avers that both individuals have a conception of the good which, according to the definition of a conception of the good given above, means that they have both chosen autonomously (rationally). He then goes on to say that the beer-drinking TV-viewer is "less able to describe" his conception but this, evidently, does not mean that he has not articulated his preference or that he has not established some set of priorities. What Dworkin means is that his preferences and the lifestyle that satisfies them are rather uncomplicated. In fact Dworkin, like Rawls, claims that where justice (or foundationalist morality) is concerned, the two conceptions are on a par. They are both regarded as preference-satisfying lifestyles.

To explain how this position can be maintained, that is, how conceptions of the good embedded in ideals can be regarded as preference-satisfying conceptions, it is important to incorporate into the neo-Kantian picture of autonomous choice-making the concept of a "second-order volition", a concept which is introduced by Harry G. Frankfurt in "Freedom of the Will and the Concept of the Person".<sup>7</sup> According to Frankfurt, every individual has a will, that is, he can form desires that move him effectively to action. He can also form "second-order desires", i.e. he can want to have certain desires. The

autonomous individual, however, can also want to have second-order desires that will move him **effectively** to action. These desires Frankfurt calls "second-order volitions" and considers the ability to form them essential to being a person. Clearly, the capacity to form rational plans of life presupposes this particular ability.

To illustrate this point, Frankfurt contrasts the autonomous drug addict with the non-autonomous one. The latter has two conflicting desires: to take the drug and to refrain from taking it (which is a second-order desire). He, however, simply weighs the two desires. He has "no stake in the conflict and can, therefore, neither lose nor win the struggle."<sup>8</sup> The autonomous addict, by contrast, "makes one of the conflicting desires his own and in so doing he withdraws himself from the other."<sup>9</sup> It is this identification with the second-order desire that turns it into a second-order **volition**.<sup>b</sup>

The point to be made here is that individuals can form the desire to appreciate or internalise systems of values that discriminate among preference-satisfactions. This, of course, involves suppressing rival desires or preferences. The scholar in Dworkin's example may have striven against his desire for beer-drinking cum TV-watching before he was able to enjoy the life of contemplation. But, even when one is moved by an ideal, one can form the second-order volition to stop being

---

<sup>b</sup>It does in no way follow from all this that the autonomous addict will actually be successful in sloughing off his addiction or that he will not need help from others to achieve this end.

affected by it. As I mentioned in the previous section, Rawls himself makes a similar point. He believes that one can easily change one's preferences and that, therefore, the adoption of a conception of the good is not necessarily dictated by them. This indicates that Frankfurt's view is not only consonant with this aspect of Rawlsian essentialism, but that it is actually a more thorough formulation of it.

Both questions raised above have been answered. Ideals, like other conceptions of the good, are chosen because they satisfy particular desires. But autonomous individuals are not compelled by their desires (of whatever order) to act in particular ways. They have the essential ability to identify and reflect on them and weigh one against the other before they go on to devise a rational plan of action.

It follows from all this that no particular conception of the good is essential to being a person. Regarding some conceptions as more valuable than others, would be making them part of the essentialist account which is the foundation of justice. However, this is not to say that all conceptions of the good are valued equally in every area of human interaction. Since the foundationalist morality is to govern only the institutional structure of society, there are various other contexts in which the evaluation and categorization of conceptions of the good are inevitable. The criteria by which conceptions of the good are measured for value are, of course, contingent.

It might be argued, however, that certain criteria for the evaluation of conceptions of the good can be found within

the neo-Kantian foundationalist position itself. I shall mention here three such standards of evaluation. The first one has already been mentioned and pertains to Rawls' so-called "Aristotelian principle" which is one of the "laws of human psychology". According to this principle, "...a person prefers the [activity] that calls upon the greater number of more subtle and intricate discriminations."<sup>10</sup> If we apply this to Dworkin's example, we arrive at the conclusion that the life of contemplation is clearly superior to the indolent life of TV-watching and beer-bibbing. Does this mean that justice requires the unequal distribution of rights to the adherents of the two differing conceptions?

Unlike the two classifications that follow, this one poses no serious problems. What Rawls calls "laws of human psychology" do not have the same status as the central aspects of his essentialist account (i.e. autonomy, personal distinctness, plurality and the desire for primary goods). The Aristotelian principle is not, therefore, a moral criterion--a dictate of foundationalist morality--but part of the effort to chart the process of rational choice-making. Choosing more elaborate or complex activities is not, in other words, essential to being a person and, therefore, the apparent inferiority of beer-drinking and TV-watching to the life of contemplation is neither a derivative nor a concern of justice.

A second classification can be made by employing a feature of Kantian moral philosophy, namely the categorical imperative.

Conceptions of the good that are adopted by the individual, not only in order to satisfy his personal preferences, but because he wills them to be adopted by everyone, could be regarded as having a special status, as being superior to preference-satisfying conceptions. Clearly the implication here is that preferences should be universalized. Although the Kantian nature of this view is unmistakable, the neo-Kantian foundationalist does not endorse it. Instead, his response is that this view is incompatible with the neo-Kantian essentialist account and more particularly with the notion of pluralism.

A third possible foundationalist classification of conceptions of the good could, perhaps, be made on a wholly different basis. Conceptions of the good are placed into two categories, the first of which includes conceptions that allow one to revise one's initial choice and make further autonomous choices, whereas the second contains those conceptions that preclude further choices and compromise one's autonomous nature. This is a distinction between "autonomous" and "non-autonomous" conceptions of the good.<sup>11</sup> To choose a conception of the latter kind, is to inflict harm upon oneself, to consciously relinquish one's essence of personhood. If this classification can be made, then does justice allow the choice of non-autonomous conceptions, or does it require the introduction of "morals laws" prohibiting their adoption?<sup>c</sup> In

---

<sup>c</sup>"Morals Laws" are defined as "laws forbidding certain powerfully seductive and corrupting vices [and which] can help one to establish and preserve a virtuous character"; Robert George, Making Men Moral (Clarendon Press: Oxford, 1983), p. 14. These can either be interpretivist principles or

order to answer this question, I will introduce David Richards' formulation of neo-Kantian foundationalism.

## 4.2 Vices as Conceptions of the Good

David Richards is, perhaps, the only exponent of a state constraint position who does not dilute the foundationalist character of his work in order to make it more palatable to his audience. As I have explained, Rawls' foundationalism in A Theory of Justice is only latent and often combined with an interpretivist approach which Rawls explicitly endorses in his more recent writings. Richards, whose doctoral thesis was supervised by Rawls, never revised his neo-Kantian position, but undertook the exploration of its implications. In his Sex, Drugs, Death and the Law, he focuses on conceptions of the good that are widely regarded (at least in Western societies) as vices, and argues, along strict foundationalist lines, that their prohibition is contrary to the dictates of justice.

Richards' "principle of (equal) respect for autonomy (or personhood)" is none other than the neo-Kantian principle of state constraint.<sup>12</sup> Richards argues that institutional arrangements governed by principles that derive from contingent moralities (or what he calls "perfectionist ideals")--whether or not they treat autonomy as among human goods protected by rights to, say, freedom of speech, religion and assembly--are unacceptable since they may compromise the autonomous nature of the individual. As he puts it,

---

principles derived from an essentialist account that is different from the neo-Kantian.

[t]he idea of 'human rights' respects this capacity of persons for rational autonomy, their capacity to be, in Kant's memorable phrase, free and rational sovereigns in the kingdom of ends. Kant characterised this ultimate normative respect for the reasonable choice of ends as the **dignity of autonomy**, in contrast to the heteronomous, lower-order ends (pleasure, talent) among which the person may choose...[T]he concern embodied in the idea of human rights is not with maximising the agents' pursuit of any particular lower-order ends [i.e. conceptions of the good], but rather with **respecting the higher order capacity** of the agent to exercise rational autonomy in choosing and revising his ends, **whatever they are.**<sup>13</sup>

Richards, like Rawls, expresses the view that what quells autonomy is not facing practical problems (e.g. lack of resources or knowledge) with the execution of a chosen plan, but living in a society whose basic institutional structure is governed by coercive principles that derive from contingent moralities exhorting the maximization of the "pursuit of...particular lower-order ends." At the same time, Richards explains in what sense this impairs or compromises autonomy: when the basic structure is governed by a contingent morality, individuals' liberty rights are violated. This does not affect directly their capacity to choose rationally.<sup>d</sup> Rather, it has a negative effect on the individuals' self-esteem, on the sense of dignity that they derive from being who they are. This is what Richards calls "the dignity of autonomy" and what Rawls calls "the primary good of self-respect". As has been

---

<sup>d</sup>A person whose liberty rights are violated is as likely to lose his capacity for rational choice-making as an individual who is persistently and systematically called abusive names is likely to lose his sense of identity. What is more easily lost is the sense of dignity that the individual derives from being who he is.

mentioned, according to Rawls' second foundationalist argument, self-respect is conditional on the fulfilment of the essential want for primary goods. It seems, however, that self-esteem is also built into the notion of autonomy of the first foundationalist state constraint argument and that it, therefore, is a fundamental element of the neo-Kantian foundationalist position.

There is no doubt that Richards' Sex, Drugs, Death and the Law is an application of the neo-Kantian state constraint position. Richards goes beyond the guarded, non-committal foundationalism of A Theory of Justice to give the neo-Kantian view on issues such as the legality of pornography and the use of narcotic drugs. His main objective is to present what we may perceive as vices or social problems as valid objects of rational choice-making. With respect to pornography he avers that

[it] can be seen as the unique medium of a vision of sexuality, a 'pornotopia', a view of sensual delight in the erotic celebration of the body, a concept of easy freedom without consequences, a fantasy of timelessly repetitive self-indulgence... [It] affords the alternative idea of the independent status of sexuality as a profound and shattering ecstasy.<sup>14</sup>

In the same spirit Richards gives an account of drug use and claims that

the psychological centrality of drug use for many young addicts in the United States may, from the perspective of their own circumstance, not unreasonably organize their lives and ends... [T]he moral criticism implicit in the concept of drug

abuse fails to take seriously the perspective...of the addict, often substituting competencies and aspirations rooted in the critic's own background and personal aspirations to organize a self-respecting social identity, which may only exceptionally require drug use... [E]ven psychological devotion to drugs may express not a physiological bondage, but critical interests of the person.<sup>15</sup>

Richards' intention clearly is to articulate specific, "reasonable" preferences and then show that the perceived vices are activities or lifestyles that satisfy these preferences. This qualifies them as conceptions of the good, that is, as valid objects of rational choice.

In doing this Richards emphasizes the contingency of both the vices *qua* conceptions of the good and the morals laws that prohibit adherence to these conceptions. Pornography, he claims, is not valuable to everyone nor could it be so. It can simply be regarded as a "unique medium of a vision of sexuality". Similarly, the use of drugs does not constitute an ideal way to organise one's life, but many people may certainly feel that it suits them best. The moral outlook that rejects or prohibits these conceptions is also contingent. For Richards, "there is something morally perverse in condemning drug use as intrinsic moral slavery when the very prohibition of it seems to be an arbitrary abridgement of personal freedom."<sup>16</sup> The overriding morality is clearly the foundationalist morality deriving from the neo-Kantian essentialist account.

Before I address directly the question about the existence of non-autonomous conceptions of the good and their

place in foundationalist state constraint theory, I will briefly discuss four criticisms of Richards' position which fail to grasp the significance of the fact that his argument is strictly foundationalist.

According to Robert George, Richards' is a "strategy of arguing from the value of something to a right [and, therefore, it] is perfectionist."<sup>17</sup> In George's view, this same strategy is followed by the legislators who are in favour of the prohibition of drug use, pornography and prostitution. It is, in other words, because Richards regards these particular activities as intrinsically good (as good for everyone) that he adopts the state constraint principle.

As I have already explained, Richards does not regard pornography or drug use or any of the other activities and lifestyles that he discusses as "positive moral goods", but as options available to people in specific circumstances, with "reasonable" preferences and perspectives. It is the rational choice-maker who is "respected" by the higher-order morality of justice and not the object of the choice. Only when an individual's actions are in violation of others' freedom (protected by the liberty and harm principles) is punitive or repressive state action warranted.

Eventually George admits that Richards' position can also be interpreted as an "autonomy-based understanding of moral personality".<sup>18</sup> He then makes a second point. He avers that Richards' argument leads to counter-intuitive, unpalatable results by allowing practices that most people would regard as offensive and contrary to any sense of decency and propriety.

George does not seem to be conscious of the fact that foundationalist political philosophy does not draw from the ideas and values that are prevalent in any actual society. The fact that Richards' prescriptions are deemed utterly immoral by many people does not discredit his argument. In fact George's point is reminiscent of the reaction of the cave-dwellers of the Platonic fable to the teachings of their travelled companion. However, at the same time George challenges the neo-Kantian contention that the foundationalist morality is compatible with contingent moral outlooks and that it can, therefore, regulate the actions of their adherents.

Rawls provides us with the means to respond to this criticism. In his view, we can talk of compatibility only at a higher (meta-theoretical) level of analysis. When we turn to specific societies we are bound to observe conflicts between certain conceptions of the good and the foundationalist morality. Racism and Puritanism are examples of conceptions which, were they to be predicated of the basic institutional structure of society, would lead to distributions of rights very dissimilar to the distribution stemming from Rawls' essentialist account. And yet, Rawls claims, such conceptions may be held and shared in the extra-political (i.e. private) world. A racist, for example, can choose to take residence in an area inhabited by members of his own race or refrain from socialising or working with members of other races. He can even join exclusive clubs that admit members on the basis of race or gender etc.. The puritan can lead an austere, celibate life and associate with people who share the same values. Only

the activities that violate the harm principle are prohibited and not the systems of values that may lie behind them.

But does Richards make this same move? Does he recognise that individuals are free to express their other-regarding preferences in the extra-political sphere? On first inspection, it seems that he does not. He not only criticises the distribution of rights in accordance with contingent moralities, but also the expression of contingent moral values outside the political sphere. With respect to the social treatment of prostitutes, for example, he states that

the moral condemnation of the prostitute rests on and expresses such isolation and denial, disfiguring the reasonable perception of the form sex takes in our lives, drawing sharp moralistic distinctions between the decent and the indecent when in fact there is a continuum of varying personal modes of sexual expression and fulfilment... When we extend to prostitutes concern and respect for their equality as persons we can see the source of the previous misperception. The failure to see the moral and human dignity of the lives of prostitutes is a moral failure...<sup>19</sup>

He goes on to say that drug addicts should not only be permitted to pursue their conception of the good, but also to "organize a self-respecting social identity..."<sup>20</sup> Those of a puritanical bent, we are told, should not snub or isolate prostitutes and drug addicts, but treat them as respectable individuals.

There are two possible interpretations of this point. First, it can be regarded as the same application of the categorical imperative put forward by Rawls in the case of racists. We have already seen what kinds of problems this view

entails.

Second, it could be the case that like Mill, Richards believes not only in the distribution of rights in accordance with foundationalist morality, but in a social reform, an introduction to the social sphere of a moral outlook that will vie against the Millian "tyranny of custom". Although the institutional structure does not directly play any part in this rivalry, the principles of foundationalist morality become also a lower-order morality to compete in the social sphere. I believe that this view has little to do with the foundationalist argument itself.

A third criticism focuses on Richards' notion of autonomy. Its author, J. Finnis, claims that this notion of autonomy--contrary to what Richards himself claims--is not at all Kantian because, according to Kant, "one has autonomy just insofar as one does in fact make one's choices not on the basis of one's interests, but out of respect for the demands of [the shared] morality."<sup>21</sup> In his "Legal Enforcement of Duties to Oneself: Kant versus Neo-Kantians", Finnis quotes extensively from Kant's moral writings in order to support his claim. His argument is based on the distinction between personal preferences and "reasons for action".<sup>22</sup> In his view, Kant, unlike Rawls and Richards, regards existing, shared moralities as the only providers of reasons for action. The upshot of this argument is that being spurred to action by desires and preferences is being abandoned to internal disorder, and since autonomy is about taming internal demons--as much as it is about staying free from external coercion--

the Neo-Kantian conception of autonomy is seriously flawed.

To the extent that the argument seeks to do damage simply by severing the links between state constraint foundationalism and Kantian moral philosophy, it is bound to be ineffective and superficial. When, however, Finnis links action triggered by desires with the kind of internal disarray that is the very opposite of autonomy, the neo-Kantian foundationalist, as we have seen, responds that the autonomous agent deals with internal demons by articulating his preferences, forming second-order volitions and internalizing value-structures. The latter, however, being contingent, cannot be elevated to the status of a morality that regulates public life. After all, the very essence of foundationalism is the belief that there can be a universal morality applicable to all societies irrespective of their history and culture.

George is the author of the fourth criticism of Richards' view. He claims that prescribing "duties to oneself" is an essential characteristic of every moral position. Since Richards' autonomy-based position does not include such duties, it does not qualify as a morality. Although the view that "duties to oneself" are essential to a morality is arbitrary (at least in the eyes of the foundationalist who has a different view of the moral and who claims to have knowledge of the true, coherent morality), it points to a possible inconsistency in Richards' words: If autonomy is the essence of personhood, should a person be allowed to relinquish his autonomy? Should not choices that preclude further choices be, somehow, prohibited? Don't we all have the duty to choose

autonomous conceptions of the good?

First of all, what has to be determined is whether there actually are any non-autonomous conceptions of the good. Not all philosophers accept the idea that one can rationally choose to relinquish one's capacity to choose rationally.<sup>e</sup> I will consider this view as I discuss specific examples of putative non-autonomous conceptions of the good.

My starting-point will be the case of voluntary slavery. Does the principle of constraint (and more specifically its component principle of liberty) allow one to choose to become a slave, to relinquish one's liberty and forgo any future use of it beyond this single act of choice-making? Mill's response to this question is that the liberty principle (i.e. the principle of On Liberty) "cannot require that [one] should be free not to be free. It is not freedom to be allowed to alienate [one's] freedom."<sup>23</sup> Mill's argument seems to be along the following lines: the liberty principle seeks to guarantee that each individual can choose what is good for him without being coerced by others. Given that the abdication of one's liberty precludes the possibility of further uncoerced choices, it cannot be permitted by the liberty principle.

Gerald Dworkin has argued that this line of argument is flawed and that Mill cannot proscribe voluntary slavery without contradicting his own foundationalist position. In an argument reminiscent of Richards' attack on the laws

---

<sup>e</sup>Susan Mendus, for example, states that "...certainly individual decisions within our lives may be autonomously or non-autonomously made. Nevertheless, that we may, so to speak, 'choose an autonomous lifestyle' is paradoxical," Toleration and the Limits of Liberalism, p. 106.

forbidding drug use, prostitution and pornography, Dworkin describes slavery as a "reasonable" conception of the good. He claims that there is nothing incoherent in a person saying, "I want to be the kind of person who acts at the command of others. I define myself as a slave and endorse those attitudes and preferences. My autonomy consists in being a slave."<sup>24</sup> Dworkin seems to be saying that, if we accept Mill's claim that liberty is prescribed because it protects autonomy, we have to face the fact that in this particular case autonomy simply has no need for liberty; it does not require its protection.

But what if the voluntary slave eventually comes to the conclusion that acting at the command of his master is no longer good for him? Surely, this is not an impossible development. The autonomy of the person who relinquished his liberty is not impaired as long as he wills to obey his master's commands. His self-respect (his "dignity of autonomy") is intact as long as he thinks that what he does is good for him.<sup>f</sup> His liberty (if he had not relinquished it) would protect his autonomy only at the point where he changed his mind about his wanting to be a slave and his master refused to release him on the grounds that he was bound by contract to remain always his slave.<sup>g</sup> It is, therefore, at this point that he "needs" his right to liberty and not during

---

<sup>f</sup>For example, professional soldiers do not lose their self-respect by obeying orders. In fact they take pride in doing so. It would be absurd to claim that professional soldiering somehow compromises one's autonomy.

<sup>g</sup>In fact one could argue that it is only at this point that he becomes a slave.

his term as a voluntary slave. It is at this point that the liberty principle becomes relevant.

David Archard's interpretation of the Millian view on voluntary slavery is compatible with this view. According to Archard, Mill suggests not that society should attach penalties to the signing of a slavery contract, but that such a contract should not be legally enforced.<sup>25</sup> Strictly speaking, Mill does not say that we have a duty not to abdicate our liberty. Rather, his point is that if the liberty principle governs the basic social structure, the idea of choosing slavery as a conception of the good is absurd.

The second case of a commonly reputed non-autonomous conception of the good is the lifestyle of the drug user. As in the case of the voluntary slave, the prospective drug user's decision to take narcotic drugs has to be respected as an autonomous choice. As Harry Frankfurt explains it is possible for a drug user to retain his capacity for autonomous choice-making even when addicted to drugs (see previous section). It is, however, possible (and some would say, probable) that some time after his initial decision the user falls in a stupor which, given his physical dependency on the drug, leaves him only one choice: to keep taking the drug. The basic difference between his case and that of the voluntary slave is that his autonomy can be compromised by internal disorder and not by external coercion. The autonomous addict, who has formed the second order volition to slough off his addiction, can easily make his intentions known to others and seek their help. But the autonomous addict, who sticks by his

original decision to take drugs, is virtually indistinguishable from the non-autonomous one who keeps taking them for purely physical reasons. The neo-Kantian foundationalist cannot, therefore, proscribe the use of drugs as a non-autonomous conception of the good.

The last putative non-autonomous conception of the good that deserves mention is suicide/euthanasia. Unlike the voluntary slave and the drug user, the person who puts an end to his own life will certainly be unable to re-evaluate his position and revise his decision. It, therefore, appears that if there is any non-autonomous conception of the good that a person could choose, it is death.

But is this view compatible with the foundationalist state constraint position? A neo-Kantian foundationalist could very well argue that the person who deliberately takes his own life is not choosing not to be autonomous. What he chooses is to cease to exist and this can hardly fit into the notion of a conception of the good. It, therefore, follows that the principle of constraint does not hold in this particular case. As Archard puts it, "suicide brings about a situation in which the 'liberty principle' no longer applies."<sup>26</sup>

Straightforward though it may appear, this point fails to answer an important question: does the foundationalist morality allow individuals to leave its jurisdiction? The foundationalist's most fundamental assumption is that public life should be regulated by a morality that protects (and in certain cases this only means "respects") what is essentially human. Given that being alive is implicit in the essentialist

account, why should the act of suicide be permissible?

A possible answer is that the individual who chooses death has first lost the dignity of autonomy (e.g. through incurable physical pain, chronic depression or dishonour). If it is possible to lose one's dignity of autonomy in a just society (and I believe most neo-Kantians would find that it is), then it is possible to lose the willingness to act autonomously and to seek death as the only option. If this is so, then justice can neither detect nor prevent this from happening. This point is alluded to by Archard himself who avers that "it could...be rational for individuals to end their life...because of [their liberty's] value being outweighed by the evils that such individuals must unavoidably suffer."<sup>27</sup> The evils Archard implies here are not the evils of injustice, but the pain and disabilities of illness and the internal disorder which they bring about.

Although there is no agreement among neo-Kantian foundationalists on whether this view is compatible with the two main constraint arguments, there is consensus on one very important point: the constraint principle does not classify any conceptions of the good as non-autonomous.

### **4.3 Culture and Conceptions of the Good**

I shall now turn to a different kind of conception of the good, one that is much broader in scope than any other discussed so far. Conceptions of the good of this kind are commonly referred to as "cultures" or "cultural structures". Critics of neo-Kantian foundationalism have suggested that the

good of cultural membership (being a communally shared good) is lost to the members of a society whose basic structure is governed by the state constraint principle. Given that the principle guarantees each individual's freedom to revise, modify or change his conception of the good and that cultures are sets of values, priorities or activities that people feel obliged to perform, respect and adhere to, it appears that the citizens of the neo-Kantian state would hardly take the trouble to subscribe to them over long periods of time. This clearly means that in a society governed by the neo-Kantian morality, cultures would dwindle through lack of committed and long-term support. The crux of this view is not only that the principle of constraint discriminates against these particular conceptions of the good and fosters asocial individualism, but that the loss of cultural membership is the loss of an essentially human attribute.

The most notable response to this position is offered by Will Kymlicka. Unlike Richards, whose attack on morals laws strengthens the impression that atomism is built into the neo-Kantian foundations, Kymlicka seeks to accommodate the communitarian critique by arguing that cultural membership is an indispensable element of the neo-Kantian position and that it should be made explicit in the formulation of the constraint principle. His position will, therefore, be the main focus of this section. My main point will be that Kymlicka overstates his case and contradicts his commitment to autonomy-based foundationalism. Nevertheless, the remedy can be found within his argument, and my aim is to redress it and

set it in line with the foundationalist constraint position.

Kymlicka's commitment to the autonomy-based foundationalism of Mill and Rawls is beyond doubt. In his "Liberal Individualism and Liberal Neutrality", he defends Rawlsian state constraint along foundationalist lines and claims that the state should not interfere with the pursuit of communal goods. As he puts it, "...liberals believe that people naturally form and join social relations and forums in which they come to understand and pursue the good. The state is not needed to provide that communal context and is likely to distort the normal processes of collective deliberations and cultural development".<sup>28</sup> And yet, in his Liberalism, Community and Culture, Kymlicka changes his mind on the issue and maintains that the preservation of minority cultures through the distribution of minority rights to all cultural minorities is a dictate of justice.<sup>29</sup> In his view, this is the conclusion that a neo-Kantian foundationalist is bound to reach once he acknowledges the fact that the context of autonomous choice is a particular cultural structure. As he puts it, "in deciding how to lead our lives, we do not start de novo..."<sup>30</sup> We examine ideals and lifestyles that have already been developed by others before us and we adopt them or modify them so as to fulfil our preferences and satisfy our desires. The articulation of our desires and the development of our preferences are achieved in terms of the available systems of value and the local language which, itself, is "not a neutral medium for identifying the content of certain activities."<sup>31</sup> To this, Kymlicka adds that

[t]he processes by which options and choices become significant for us are linguistic and historical processes. Whether or not a course of action has any significance for us depends on whether, or how, our language renders vivid to us the point of that activity. And in the way our language renders vivid these activities is a matter of our cultural heritage.<sup>32</sup>

Kymlicka mistakenly draws the conclusion that the individual's autonomy depends on the security of his culture (his index of choices, as it were) and that the security of the latter must, therefore, be guaranteed.<sup>33</sup> The flaw in the argument is that once autonomy is accepted as a foundation, the view that the individual is embedded in his own culture can only be held in a particular way. It cannot mean that the self is only defined by his culture or society.<sup>b</sup> To suggest so is to reject foundationalist philosophy and embrace interpretivism instead, which is something that, I assume, Kymlicka is not willing to do. His mistake, therefore, is that in trying to placate the communitarian critics of neo-Kantian foundationalism, he tried to combine two very different, irreconcilable positions: neo-Kantian foundationalism and the interpretivist thesis. This is picked up by Jeremy Waldron, who asserts that "...in developing his case, Kymlicka is guilty of something like the fallacy of composition."<sup>34</sup>

---

<sup>b</sup>When not accompanied by some account of what is essentially human, the claim that the individual is somehow "attached" to his culture becomes a meta-theoretical statement, a rejection of the universalist essentialism of foundationalism, an echo of Herder's much quoted phrase "Not a man, not a country, not a people, not the past of a people, not a state are like one another. Consequently, the true, the beautiful and the good in them are also not alike" (quotation from Christopher Berry, Human Nature, p. 69).

This is not to say that the neo-Kantian foundationalist does not appreciate the significance of cultural membership. In fact, although he cannot prescribe the protection of cultural minorities or assign rights to all of them, there are cases in which he finds that the assignment of minority rights is warranted. Furthermore, he takes the view that the state can, indeed, contribute to cultural enrichment, but without assuming a protectionist role.

I believe that it is possible to interpret Kymlicka's recent argument along these lines. Although he overshoots his target and promises his audience more than he can actually derive from his foundations, this less protectionist position is compatible with most elements of his recent argument and fully in line with the earlier one.

In his Liberalism, Community and Culture, Kymlicka asserts that the morality that should govern the basic structure of society is not to be found in any particular cultural structure or structures, but to be derived from the neo-Kantian essentialist account. In his view, cultural structures have instrumental value: they provide their members with the range of choices that is needed for the exercise of autonomous choice-making. In Kymlicka's own words, we

...should be concerned with the fate of cultural structures, not because they have some moral status of their own, but because it's only through having a rich and secure cultural structure that people can become aware, in a vivid way, of the options available to them, and intelligently examine their value.<sup>35</sup>

But what does being "concerned with the fate of cultural structures" mean to the neo-Kantian foundationalist? To answer the question we must first examine Kymlicka's use of "culture", "cultural structure", "cultural enrichment" and "conception of the good".

Apparently, Kymlicka uses the first two of these terms as synonyms. He takes them both to mean the sum of shared values, customs, habits, conventions and activities (i.e. the sum of particular conceptions of the good) of a linguistic or ethnic group. Kymlicka refers to Quebecois culture, Inuit culture, Islamic culture and even to English culture (with reference to Lord Devlin's views on the "dangers" to the English values and way of life). This means that, in Kymlicka's view, a cultural structure is a set of particular, integrated conceptions which is insulated from the outside world. To use Jeremy Waldron's phrase, a cultural structure, for Kymlicka, is a "snapshot version of it".<sup>36</sup> Accordingly, "cultural membership" means having exclusive access to an integrated set of conceptions of the good. It follows, that when Kymlicka talks about "being concerned with the fate of cultural structures", he suggests that the neo-Kantian foundationalist morality requires the protection of cultures qua integrated, sequestered sets of conceptions of the good.

Nevertheless, implicit in Kymlicka's argument is the notion of "cultural enrichment" which comports poorly with his notion of cultural structure and cultural membership. Kymlicka clearly believes that a rich cultural structure is a condition

of autonomous choice-making since it provides a variety of means by which the individual can articulate his preferences and plan their fulfilment. This means that he does not have to and, indeed, cannot accept the narrow sense of "cultural structure"; he cannot think of a culture as a stylized, ready-made lifestyle that has to be protected by means of minority rights. Instead, a culture must be understood as a structure into which individuals may bring elements of different or even rival cultures as they revise, modify and develop its values and practices through their autonomous choices.

Kymlicka must, therefore, conclude that "being concerned about cultural structures" does not mean preserving each and every one of them separately by assigning separate jurisdictions, but guaranteeing freedom of choice for the members of cultural minorities and widening the range of options by encouraging individuals to play an active role in the development of their culture.

Not surprisingly, Kymlicka has not excluded this conclusion from his mixed argument. On one occasion he states clearly that "the government [should] ensure...an adequate range of options by providing tax credits to individuals who make culture-supporting contributions in accordance with their personal perfectionist ideals."<sup>37</sup> More significantly, he rejects the view of a cultural structure as a "favourite snapshot" when he asserts that the reactionary (defensive) face of Islam is not its true or genuine expression and that the fundamentalists who claim that without restriction on the freedom of speech, press, religion, sexual practices etc. of

its own members, their culture will disintegrate, should not be taken as the true spokesmen for Islam nor as its true adherents. To this he adds that "...it is wildly implausible to suppose that allowing individuals freedom of religion or sexual practices would lead to the breakdown of that [i.e. the Islamic] community, be it in England or Iran."<sup>38</sup> This is a view of a cultural community that an uncompromising interpretivist would reject. As it stands, however, it would get the endorsement of many a neo-Kantian philosopher.<sup>i</sup>

It is, therefore, my view that despite Kymlicka's attempt to appease the communitarian critics of state constraint by ascribing minority rights, his neo-Kantian foundationalism is difficult to conceal. In fact, one could say that his argument is but a juxtaposition of the interpretivist thesis and the neo-Kantian position and that it cannot, therefore, yield his conclusions.

Now, although neo-Kantian foundationalism does not call for the distribution of cultural minority rights to each and every cultural minority within society, it does warrant the assignment of minority rights in specific circumstances. I will now argue that the Kymlicka of Liberalism, Community and Culture reaches this conclusion, but decides, rather unjustifiably, to go beyond it and prescribe minority rights

---

<sup>i</sup>Thus stated, Kymlicka's position is what Waldron calls "the cosmopolitan alternative" in his "Minority Cultures and the Cosmopolitan Alternative", in W. Kymlicka (ed.), The Rights of Minority Cultures (Oxford University Press: Oxford, 1995), pp. 93-122. It is also no different from the view expressed by Allen Buchanan in "Assessing the Communitarian Critique of Liberalism", Ethics, 99 (1989), pp. 852-882.

for all cultural minorities.

In Liberalism, Community and Culture Kymlicka actually proffers a reformulation of the second foundationalist constraint argument. As I have already explained (see 3.6), the starting-point of the second foundationalist constraint argument is that individuals qua rational choice-makers have the essential desire for primary goods. Kymlicka now claims that cultural membership is one such good. As he puts it, "[t]he primary good being recognised is the cultural community as a context of choice, not the character of the community or its traditional ways of life, which people are free to endorse or reject."<sup>39</sup> Given that primary goods are actually conditions required for the pursuit of any particular chosen goal, cultural membership, as a "context of choice", must be regarded as such a good.

This addition to the class of primary goods yields not just the Rawlsian harm and liberty principles, but what I will call the "cultural harm" principle. The argument itself, like the second foundationalist constraint argument, is quite straightforward: if, indeed, cultural membership is a primary good, a person who is "undeservedly" deprived of his cultural membership is entitled to some form of restitution. But what exactly constitutes an undeserved loss of cultural membership?

First of all it is important to understand what does not constitute the loss of cultural membership in Kymlicka's view. To do so one has to appreciate the role of the market in foundationalist constraint theory. The second foundationalist

constraint argument, as it is presented in this chapter, binds together the notions of autonomy, equality and responsibility. Individuals making autonomous choices from a position of (near) equality of primary goods are responsible for the outcome of their choices and are not eligible for compensation granted from the institutional structure. All foundationalists discussed in this chapter claim that the market, under certain conditions, is consonant with this foundationalist picture.

Kymlicka himself asserts that

...given certain background conditions, the market assesses the cost to others of my choices. Under these conditions, an efficient market distribution of resources is a fair one...[Foundationalists] value the market (or something that replicates the results of the market) not because maximizing wealth or preferences is a good itself, but because markets provide a way of measuring what is in fact equitable.<sup>40</sup>

Kymlicka not only endorses the market, but talks about the "cultural marketplace"<sup>41</sup>, where cultural membership is bought or sold at a price. Within the market individuals quite literally "shop for culture". They make choices responding to the realities of the market as well as to their own personal preferences, and in so doing, they transform their culture (e.g. in order to minimize the cost of membership), convert to rival cultures by buying membership to them or, as so often happens, by assimilating particular features (conceptions) of different cultures. One can, therefore, belong to what may be called English culture, but, at the same time, follow Oriental religious practices, read Russian literature, cook Chinese

food and collect African tribal art. In this context, the decline of a particular "snapshot version" of a culture is quite possible, and even probable.<sup>j</sup> But does it entail the loss of cultural membership? And if it does, are the members of a cultural community responsible for it and, therefore, ineligible for compensation?

Kymlicka's answer is that cultural decline, or transformation, does not involve the loss of membership at all. Members revise or move away from their culture gradually. When, in the eyes of the conservative members of the cultural group, the culture appears to be in a crisis, most members have already opted for a revised version of it or have secured membership to a neighbouring or rival culture.

For Kymlicka, French Canada is a case in point. A radical transformation of the Quebecois culture began in the 1960s. Slowly but steadily members of the local culture moved away from institutions which traditionally characterised French Canadian life (e.g. the Roman Catholic Church, the Union Nationale party, parochial schools) and in parallel the number of Anglophones grew. Membership was not lost because the fate of Quebecois culture was actually determined by the autonomous choices of its members, who freely chose a specific

---

<sup>j</sup>Of course, what declines is a particular expression of the culture which has been recorded in time and not the culture itself. For the neo-Kantian foundationalist, therefore, cultural decline is not an unintended or undesirable development. It simply is a consequence of the autonomous choices of individuals who rationally form preferences and respond to the ever-changing realities of the cultural marketplace. The neo-Kantian would, therefore, understand decline as transformation. It is only to the reactionaries within each cultural group that decline has negative connotations.

interpretation of what they perceived as their culture. Those with a nostalgic view of earlier stages in the development of the culture have no right, according to foundationalist morality, to impose their favourite version of the culture on others.

But if cultural membership is not lost with the demise of a culture within the cultural marketplace, then how is it lost and when does its loss warrant compensatory measures? Kymlicka's answer is straightforward: membership is lost with the violent expansion of rival cultures.

The destruction of the native North American cultures is a case in point. Native Americans were never given the option of revising their culture and making it viable in a cultural marketplace. Their traditions and ways of life were destroyed by the invading armies of a culture that regarded them as savages. As a result, they undeservedly lost both the good of cultural membership and that of self-respect. Since this was not the result of any individually or collectively made choice of the American Indians themselves, the neo-Kantian concludes that native Americans are entitled to a form of compensation that will bring about a situation approximating the status quo ante.

For Kymlicka, this can be achieved only through the introduction of the kind of minority rights that are presently in place in Canada and the United States. These rights aim at the protection of native Americans qua members of their violated culture and involve the establishment of "special political jurisdictions over which Indian communities have

certain guaranteed powers, and within which non-Indian Americans have restricted mobility, property and voting rights."<sup>42</sup> This unequal distribution of rights is a dictate of foundationalist morality and more specifically of a "cultural harm" principle which could be formulated in the following way:

encroachments made upon a cultural community outside the cultural marketplace warrant the issuing of minority rights that compensate its members for losing the primary good of cultural membership.

To attribute this argument to Kymlicka is to say that, despite his official position in favour of group rights for all cultural minorities, he does not understand personal choices to be confined within a single culture. The fact that individuals revise and develop their own cultural structure does not mean that they choose only between the different conceptions of the good that can be found within it. Instead, they choose features of other cultures and seek to incorporate them into their own. In the end, the cultural borders may become virtually indiscernible through a process of what Salman Rushdie calls "change-by fusion, change-by-conjoining".<sup>43</sup> The outcome of this process does not concern the neo-Kantian foundationalist.

In my view, therefore, the neo-Kantian Kymlicka goes no further than the conditional assignment of minority rights and this makes his position defensible against John Danley's criticism.

In his article "Liberalism, Aboriginal Rights and

Cultural Minorities", Danley claims that Kymlicka fails to show the significance of culture in liberal philosophy and politics and asserts that Mill's On Liberty is more successful, in that respect, than Kymlicka's book. Danley clearly fails to discern Kymlicka's reformulation of the second foundationalist constraint argument and concludes that "for Kymlicka the relevant difference [between native American cultures and other ethnic cultures] is that the aboriginal cultures are more vulnerable and hence require more drastic measures for their protection."<sup>44</sup> But if aboriginal cultures are vulnerable, he says, so are the Italian, German, Irish and other ethnic cultures of North America. And if indeed minority rights are awarded to all vulnerable cultural minorities, the result will be the "Balkanization" of the United States, the creation of "a crazy quilt of different jurisdictions and different languages, a jumble of different sets of rights."<sup>45</sup>

This would have been the right conclusion, had Kymlicka not made it clear that cultures, according to neo-Kantian foundationalism, are more resilient than the conservatives within each cultural community tend to think. The line separating Indian from Western European ethnic cultures is not drawn arbitrarily. European immigrants to the Americas could not have hoped to maintain their cultural heritage and pass it on intact to the next generation as if they had never left their native lands. They knowingly and willingly entered the cultural marketplace in the new countries.<sup>k</sup> The Indians, on

---

<sup>k</sup>In his discussion on the claims and grievances of the Pakistani minority in Britain, Bikhu Parekh regards the view that by emigrating to Britain Pakistanis consented to its way

the other hand, were given no choice. After invasion and denigration, the remnants of their communities could not possibly have been expected to compete in the cultural marketplace. In fact, the very rules of the market were never in place where they were concerned.

Of course, Danley's understanding of Kymlicka's position is understandable for, as has been mentioned, Kymlicka overshoots his target by endorsing the "embeddedness thesis" which is irreconcilable with the autonomy-based element in his position. Nevertheless, the main point to be made here is that the notions of culture and cultural development are not alien to neo-Kantian foundationalist philosophy. Individuals transform the cultural structures within which they are brought up by forming the second-order desire to experiment with, adopt and pursue conceptions of the good which are external to their culture. By exploring this idea, the neo-Kantian Kymlicka broadens the scope of foundationalist state constraint.

---

of life and incurred an obligation to abide by its laws, norms and values, as totally inaccurate and misleading. He states that Britain recruited Pakistanis "to help re-build its post-war economy in full knowledge of who they were and what they stood for." He, therefore, concludes that it is Britain who is under an obligation to respect and protect Islamic culture. (Bikhu Parekh, "The Rushdie Affair: Research Agenda for Political Philosophy", in Will Kymlicka, The Rights of Minority Cultures, p. 310). This kind of argument bears little relevance to the neo-Kantian position expressed here, since the British state is not governed by the constraint principle (for instance, it has a blasphemy law applicable only to Christian belief). For the neo-Kantian, it is not a particular way of life that the newcomers are expected to conform to, but the higher-order morality that derives from an account of what is essentially human. By doing so they are bound to be more exposed to the influence of other cultures than they would be in a single-culture society.

## Notes

1. John Rawls, A Theory of Justice (Oxford University Press: Oxford, 1992), p. 551.
2. Ibid., p. 551.
3. Ibid., p. 551.
4. Ibid., p. 552.
5. Brian Barry, Political Argument (Harvester & Wheatsheaf: London, 1990), p. xlv.
6. Ronald Dworkin, "Liberalism", in A Matter of Principle (Cambridge University Press: Cambridge, 1985), p. 191.
7. Harry Frankfurt, "Freedom of the Will and the Concept of the Person" in Garry Watson (ed) Free Will (Oxford University Press: Oxford, 1982), pp. 96-110.
8. Ibid., p. 104.
9. Ibid., p. 105.
10. John Rawls, supra note 1, p. 412.
11. For instance, see Mendus, Toleration and the Limits of Liberalism (Macmillan: London, 1989), p.106.
12. David Richards, Sex, Drugs, Death and the Law (Rowman: Totowa, 1982), p. 18.
13. Ibid., p. 9.
14. Ibid., p. 135. (Emphasis mine).
15. Ibid., pp. 176-177. (Emphasis mine).
16. Ibid., p. 177.
17. Robert George, Making Men Moral (Oxford University Press: Oxford, 1993), p. 144.
18. Ibid., p. 146.
19. David Richards, supra note 12, pp. 126-127.
20. Ibid., p. 177.
21. John Finnis, "Legal Enforcement of 'Duties to Oneself': Kant Versus Neo-Kantians", Columbia Law Review, 87 (1987) , p. 441.

22. Ibid., pp. 144-145.

23. J. S. Mill, On Liberty (Penguin: Harmondsworth, 1979), 173.

24. Gerald Dworkin, "Paternalism: Some Second Thoughts", in R. Sartorius (ed.), Paternalism (University of Minnesota Press: Minneapolis, 1983), p. 111.

25. David Archard, "Freedom Not to Be Free: The Case of the Slavery Contract in J.S. Mill's On Liberty" in The Philosophical Quarterly, 40 (1990), pp. 454-455.

26. Ibid., p. 462.

27. Ibid., p. 462.

28. Will Kymlicka, "Liberal Individualism and Liberal Neutrality", Ethics, 99 (1989), p 904.

29. Will Kymlicka, Liberalism, Community and Culture (Clarendon Press: Oxford, 1991). See also Kymlicka, Multicultural Citizenship (Oxford University Press: Oxford, 1995).

30. Will Kymlicka, Liberalism Community and Culture, p. 164.

31. Ibid., p. 175.

32. Ibid., p. 165.

33. Ibid., p. 169.

34. Jeremy Waldron, "Minority Cultures and the Cosmopolitan Alternative", in Will Kymlicka (ed.), The Rights of Minority Cultures (Oxford University Press: Oxford, 1995), p. 106.

35. Will Kymlicka, Liberalism Community and Culture, p. 165.

36. Jeremy Waldron, supra note 34, p. 110.

37. Will Kymlicka, "Liberal Individualism and Liberal Neutrality", Ethics, 99 (1989), p. 895.

38. Ibid., p. 168.

39. Ibid., p. 172.

40. Ibid., p. 185.

41. Will Kymlicka, supra note 37, pp. 883-905.

42. Will Kymlicka, Liberalism Community and Culture, p. 136.

43. Quoted in Waldron, supra note 34, p. 93.

44. John Danley, "Liberalism, Aboriginal Rights and Cultural Minorities", Philosophy and Public Affairs, 1992, p. 169.

45. Ibid., p. 180.

## CHAPTER 5

### Interpretivist State Constraint

Having presented state constraint as a foundationalist position, I now turn to interpretivist state constraint, i.e. the view that the state constraint principle is latent in the political practices of Western democratic societies, or, to put it differently, that it depends for its formulation and justification on the specifically political ideas that are prevalent in Western democracies. I specifically focus on two interpretivist arguments: first, what I call the "direct" argument and, second, the Rawlsian constructivist argument. In the former the interpretivist philosopher presents what he regards as widely shared ideas and general expectations, as well as historical and sociological facts, as evidence in support of the view that (his formulation of) the state constraint principle is latent in Western political practices. I attribute this argument to Rawls despite the fact that he never puts it forward in a systematic fashion. My reason for making this move is the following: I believe that Rawls' official, constructivist argument is offered as an improvement on the direct argument but is, as I will argue in 5.2, suspect from a meta-theoretical point of view. Rawls' interpretivist state constraint position is, therefore, safer only in the form of a direct argument similar to the one I piece together

drawing from A Theory of Justice and Rawls' later writings.

The constructivist argument, presented in the second part of the first section, is far more elaborate than the direct argument and incorporates various levels of interpretation. Rawls argues that members of democratic societies share certain moral views and self-perceptions which include the idea of the "moral person". He then constructs a hypothetical process, the Original Position, which incorporates his interpretivist findings as well as what he regards as two important elements of the democratic process: representation and proceduralism. In the Original Position the fictitious representatives of "actual persons" assemble to decide on principles that would govern the basic structure of a democratic society. One of these principles is that of state constraint (or, to put it differently, the liberty and harm principles).

The constructivist argument differs from the direct argument in that it is a system of interpretive findings interacting with one another in accordance with a process which seems to be characteristic of Western democratic societies.

Now, portraying Rawls as an interpretivist philosopher is not putting forward a far-fetched interpretation of his philosophy. In recent writings Rawls explicitly endorses interpretivism when he states that "the aim of political philosophy, when it presents itself within the political culture of a democratic society, is to present it [i.e. the democratic society] with certain conceptions and principles

congenial to its most essential<sup>a</sup> convictions and historical traditions."<sup>1</sup> Rawls now avers that the principles that he suggests should govern the basic structure of society are characteristic of contemporary Western political culture, a culture that comprises "the main institutions, the historical traditions [and] the shared fund of implicitly recognized basic ideas and principles."<sup>2</sup>

As has been mentioned, although many of the arguments in A Theory of Justice are interpretivist in character, the Rawls of A Theory of Justice does not seem to be committed to any one view of political philosophizing. The recent, self-conscious move towards interpretivism aims at satisfying those critics who having read Rawls' book as a foundationalist treatise, accused its author of showing disregard for the on-going, lived-in character of Western social and political tradition.<sup>b</sup> However, for some commentators, this move is also a self-critique. Rawls accepts the interpretivist critique based on the foundationalist reading of his book, endorses interpretivist meta-theory and offers a restatement which is

---

<sup>a</sup>The use of the word "essential" in the context of an interpretivist argument is ill-advised. Although the interpretivist studies the historical development of moral and political practices, ideas and self-perceptions and the circumstances under which they became prevalent, he does not believe that there are any permanent, immutable features that are unique to the society he investigates. For him, a tradition is under constant (however slow or subtle) change. Rawls is not always conscious of this point. It may, however, be the case that by "essential" he simply means prominent.

<sup>b</sup>Here I have in mind Michael Walzer, Charles Taylor, especially in Sources of the Self (Cambridge University Press: Cambridge, 1990), and Michael Sandel in Liberalism and the Limits of Justice (Cambridge University Press: Cambridge, 1982).

meant to be free of all foundationalist (universalist) elements.<sup>3</sup>

It appears, therefore, that in the light of these recent developments it does make sense to attribute to Rawls both a foundationalist and an interpretivist state constraint principle. The latter is, of course, bound to be formulated and justified in a different manner. Interpretivist state constraint must be shown to be characteristic of Western (American) political culture, to be tacitly recognized or already (partially) institutionalised and to be consonant with shared meanings and self-perceptions of the members of modern democratic societies. I will now present the direct constraint argument as a Rawlsian position.

### 5.1 The Direct Interpretivist Constraint Argument

From a methodological point of view, the direct argument is quite straightforward. The interpretivist philosopher presents evidence intended to prove that a specific principle of justice is latent in the public morality of the particular society he investigates. The argument of Michael Walzer's Spheres of Justice, for instance, is such an argument. Walzer defends principles of justice which, first, circumscribe different "distributive spheres" or "domains" and, second, dictate how goods should be distributed within each domain.<sup>4</sup>

---

<sup>3</sup>It should be made clear that the spheres are not related to the explanationist's philosophical conceptions (understood sub specie aeternitatis). For the interpretivist a "sphere" is the scope (or jurisdiction) of a prescribed principle rather than a conceptual whole (e.g. the Oakeshottian civitas) comprising essential features (or postulates).

Throughout his book Walzer maintains that the proposed principles are latent in Western (American and European) public moralities. He presents historical, sociological and even anthropological evidence to prove his point. He focuses on the history of the development of social institutions but also draws material from literature and mythology. Although Rawls offers only fragments of a direct interpretivist argument he endorses Walzerian methodology when he declares that "[t]he reasons...I give for my view are historical and sociological... What...I hold is that we must draw the obvious lessons of our political history since the Reformation and the Wars of Religion, and the development of modern constitutional democracies."<sup>4</sup>

Of course, the similarity with Walzer is not strictly methodological. Like Walzerian principles of justice, the Rawlsian principle of state constraint can be seen both as a distributive principle (governing the distribution of liberties to citizens) and as a principle marking the limit between two separate distributive domains: the political and the social or extra-political. I would, therefore, say that there is a distinct Walzerian flavour in the direct interpretivist constraint argument, whereas the constructivist argument, as we shall see later on, is an attempt on Rawls' part to produce an interpretivist style of his own and at the same time to deal with the main weakness of interpretivist philosophy, namely the lack of fixed, external criteria of evaluation.

As we saw in the previous chapter, neo-Kantian

foundationalist morality does not dictate the prohibition or restriction of any conceptions of the good. The foundationally derived state constraint principle grants everyone the "greatest equal liberty compatible with a similar liberty for others" and allows all individuals to adopt, develop, and pursue the conceptions of the good that they autonomously choose. Only "modes of conduct" which interfere with the liberty of others ought to be restricted or prohibited. The crux of the argument in support of this principle is that prohibitions against the adoption and pursuit of conceptions of the good compromise the autonomous, essential character of those who choose them.

The formulation/justification of the interpretivist state constraint principle (and the component principles of liberty and harm) is quite different. This time "liberty" is not taken in the gross and regarded as a general idea meaning the absence of moral (and physical) coercion, but a list of basic liberties drawn from the political culture of modern democratic societies. In Rawls' words, the interpretivist philosopher must show "that the scheme of basic liberties as a family is part of a coherent and workable conception of a democratic regime and, moreover, a conception that is congruent with its most essential<sup>[4]</sup> convictions."<sup>5</sup>

Rawls' list of basic liberties include "political liberty (the right to vote and be eligible for public office) together with freedom of speech and assembly; liberty of conscience and

---

<sup>4</sup>Once again, it is important to emphasize that "essential" in the context of interpretivist philosophy means prevalent or widely shared.

freedom of thought; freedom of the person along with the right to hold (personal) property; and freedom from arbitrary arrest and seizure as defined by the concept of the rule of law."<sup>6</sup> These are presented as liberties that make sense to the members of democratic societies; they are features of a living political tradition with its own "thick" moral/political vocabulary. It can, therefore, be said that there is a tacit consensus on the relevance and significance of these liberties in modern democracies. With regard to the freedom of speech, for example, Rawls states that "within our tradition there has been a consensus that the discussion of general, political, religious and philosophical doctrines can never be censored."<sup>7</sup>

The interpretivist liberty principle, therefore, takes the following form:

Each person has an equal right to a fully adequate **scheme of equal liberties** which is compatible with a similar scheme of liberties for all.<sup>8</sup>

The harm principle is, of course, adapted to this formulation of the principle. Justice now requires the prohibition, restriction and punishment of modes of conduct which interfere with the basic liberties and non-injuries of others.

What is not clear in all of this is why the basic liberties are individuated (i.e. distributed to individuals) and distributed equally. The answer is that, in terms of the direct argument, individuation and equal distribution are simply included in the shared meaning of liberty in democratic societies. Liberty as a "scheme of basic liberties" is equally

allotted to every citizen because the members of democratic societies traditionally understand it or tacitly acknowledge it to be so. The alternative interpretivist way to justify the equal distribution of liberty is to show that citizens actually understand themselves as equals and that this notion of equality somehow translates into an equal distribution of the basic liberties. It is this alternative route that Rawls takes in his constructivist argument and which I will examine in the next section.

Now, as Rawls avers, "basic liberties...take certain questions off the political agenda."<sup>9</sup> What this means is that the constraint principle can be presented not only as a liberty/harm principle (in which case emphasis is given to its distributive function) but as a principle that, using the terminology of Michael Walzer, guards the limits of distinct distributive domains. As has been mentioned, Rawls finds that there are only two such domains, the political and the extra-political. The distribution of goods in the former cannot be based on some evaluation of conflicting conceptions of the good because that would violate the individuals' basic liberties. Questions about the cogency of (foundationalist) philosophical positions, about the existence of God and the propriety of lifestyles and principles of personal virtue are, therefore, left out of the political agenda. Hence, state constraint is presented as an exclusionary principle that reads as follows:

[T]here is no political...evaluation of conceptions of the good within the limits permitted by justice.<sup>10</sup>

It should be emphasized that the state constraint principle (in either of the two formulations) is not derived from some position of moral/religious scepticism or any other conception of the good. Rather, the argument is that in Western democracies conceptions of the good are not perceived as political principles that could govern the distribution of what the state is tacitly expected to distribute--namely liberties, powers, wealth, income and opportunities. With respect to this point Rawls avers in A Theory of Justice that in Western democratic societies conceptions of the good are tacitly recognized to be "imprecise as political principles and their application to public questions...to be unsettled and idiosyncratic."<sup>11</sup> The fact that the main political forces in Western democracies are "catch-all" political parties rather than sectarian organizations can be offered as evidence in support of this point. What has been branded as "sectarian politics" (e.g. in Northern Ireland) should be regarded as an historical aberration and the interpretivist cannot draw from there valid **political principles**. This, of course, does not mean that the members of sectarian groups are not entitled to the scheme of basic liberties mentioned above. It only means that the liberty and harm principles (along with the Rawlsian second principle of justice) are derived not from any particular religious and philosophical doctrines but from a specifically political conception governing the distribution

of particular goods which are widely regarded as necessary to the adherents of all conceptions of the good. All such conceptions administer non-political goods which are prescribed by philosophical truth, religious belief and conceptions of personal virtue or character. Hence, the distinctly political conception is thought to have a limited range (or jurisdiction) and to lack the authority to evaluate the differing, non-political conceptions of the good. In Rawls' own words, "...there are no resources within the political view to judge conflicting conceptions [of the good]."<sup>12</sup>

It clearly follows that when state officials evaluate conceptions of the good, they contravene the principles implicit in the public morality and operate outside the political domain. They take sides in the disagreement between the advocates of different conceptions of the good and practically enforce (or help to enforce) one or some of these conceptions either by means of propaganda, or by distributing goods (like divine grace) which are not perceived to be suitable for state distribution, or by distributing the proper goods, but with regard to individuals' religious or intellectual partisanship. In all these cases there is interference with the basic liberties of individuals.

Rawls offers the following interpretivist argument in support of this view. First, he notes the fact of pluralism and avers that Western societies harbour a "...diversity of religious, moral, intellectual and philosophical doctrines, which adhere to conflicting ideals of personal virtue which

are to govern all parts of one's life."<sup>13</sup> He then states that the fact of pluralism can only be altered by "the oppressive use of state power," and maintains that such use of state power is not congruent with the specifically political conception which is prevalent in Western democracies. One could also add to this that the actual institutional structure of democratic societies maintains and even enhances pluralism and that, therefore, the principle seems to be already institutionalized.

Conceptions of the good are, of course, evaluated in the non-political domain. Within religious sects, academic circles, cultural groups, clubs and associations theories are scrutinized, lifestyles criticized, moral outlooks taken on faith or rejected out of hand. Social organizations often link the distribution of non-political goods (e.g. divine grace, knowledge of philosophical truths, physical fitness, inner peace, companionship) with their members' acceptance of particular values and compliance with specific directions. Each member is, however, at the same time, a citizen and as such has an equal right to the scheme of basic liberties. These protect him both against state interference and against harm inflicted to him in the non-political domain.<sup>14</sup>

---

<sup>13</sup>For instance a religious group cannot deprive its members of any of the liberties mentioned above. The threat of expulsion from the group is the usual means of inducing members to fall into line but it does not constitute harm. As long as the option to leave is open to all members, the principle is not violated. Of course, in many Western societies it is often the case that religious or ethnic groups seek to punish or reform members who have sought to distance themselves from them. For instance certain Muslim communities in Britain exercise pressure to those (especially the women) who have opted for lifestyles regarded as incompatible with

According to Rawls, the fact that we can distinguish between the political and the non-political and that we obtain principles of justice from the former does not mean that "...the participation by everyone in democratic self-government is regarded as a pre-eminent good for fully autonomous citizens."<sup>14</sup> He avers that the view that political life should have a central place in Western societies "...is but one conception of the good among others" and is derived from a particular (Aristotelian) philosophy. Rawls' point is, I believe, congruent with his interpretivism. The growing political apathy among citizens in the United States does not prove that the specifically political conception (implicit in which is the state constraint principle) is no longer a feature of Western public morality.

A more thorough formulation of this argument would be the following: within the political domain there is disagreement as to "...how more exactly to draw the basic liberties when they conflict...; how to further interpret the requirements of distributive justice and finally...questions of policy..."<sup>15</sup> Traditionally, disagreements are resolved through reasoned

---

their religious dogmas. This, of course, is in violation of the principle. A more controversial case is that of prominent members (e.g. authors and artists) who, without always exiting, try to revise the group's conception of the good and to introduce changes that may weaken the group's cohesion or change its structure. Being distinguished, influential people, they are thought by the conservative elements to be a threat to the group's existence even after they have been expelled. Thus leaders of Muslim communities in France have issued the fatwa against prominent Muslim scholars and in Britain the Rushdie affair kindled a militant tendency within the Muslim minority. But clearly the revisionists do not commit harm since they do not violate the liberties (as these are understood in Western democracies) of their co-religionists and any attempt to silence them constitutes harm.

discussion (in which the proponents of different distributive proposals and policies state their case in an effort to convince others and gain their support) and voting (which determines which proposal has the widest support). The arguments presented in support of different distributive plans during this process do not connect the proposed distribution with a particular conception of the good. They do not, in other words, present the distribution as a dictate of a conception of the good (a religious, philosophical or intellectual doctrine) or as a prerequisite for the realization of such a conception. Rather, the arguments in favour of or against any proposals are presented in terms of the specifically political conception which is characteristic of Western democratic culture. This is not related with the levels of political participation (the voting and debating) in these societies. The fact that participation dwindles does not mean that the distinction between the political and the non-political is no longer acknowledged by the members of Western societies. It could mean that unresolved, political issues concern a diminishing number of citizens, but not that the political conception comprising the principles of justice is no longer prevalent.<sup>f</sup>

This discussion of Rawls' direct interpretivist argument

---

<sup>f</sup>In "The Idea of an Overlapping Consensus", p. 14, Rawls suggests that political participation and the development of "political virtues" are of instrumental value since they can help guard against abuses of power (i.e. violations of the state constraint principle). Clearly this is a political (as opposed to philosophical-interpretivist) consideration and is in no way incompatible with the interpretivist state constraint position.

would be incomplete without mention of the notion of "public reason". In Rawls' view, the distinctly political conception includes, along with the principles of justice, "...certain guidelines of inquiry and publicly recognized rules of assessing evidence" governing the application of the principles of justice. More specifically, it includes "...the shared methods of, and the public knowledge available to, common sense, and the procedures and conclusions of science" which are not controversial.<sup>16</sup> Implicit here is the view that although in Western societies there are more than one scientific perspective and one notion of common sense (all of them undergoing changes over time), the differences between them are not deep enough to suggest that there is no single one, distinctly Western, notion of public reason.

Before I conclude this discussion of the direct state constraint argument, I should touch on the subject of the "priority of liberty" over the other goods distributed through the basic institutional structure. Over the years Rawls has offered different arguments in support of this view. In A Theory of Justice, for example, the priority of liberty is based on the disputable claim that beyond a fairly minimal point, the strength of the preference that people have for an increase in liberty diminishes less rapidly than that which they have for an increase in wealth. One could argue that this universalist claim is better suited to the foundationalist argument discussed in Chapter 3 in the sense that it can more easily be incorporated into an essentialist account rather than be presented as a characteristic of the members of

Western democratic societies. I will not, therefore, examine the validity of this claim within the context of interpretivist state constraint.

In "The Basic Liberties and Their Priority" Rawls proffers two further arguments which picture the priority of liberty as an essential aspect of the democratic political tradition. One of these fits into the direct interpretivist approach whereas the other is better suited to Rawlsian constructivism and will, therefore, be discussed in the next section.

In Rawls' view, constitutional history "suggests that principles to regulate economic and social inequalities, and other distributive principles, are generally not suitable as constitutional restrictions."<sup>17</sup> By contrast the liberty/harm principle has been regarded as a suitable constitutional restriction and has been inscribed in various written constitutions. For Rawls this is an indication that the priority of liberty over the other public goods is characteristic of Western political morality.

But is all this an accurate "reading" of Western democratic society? Does Rawls interpret correctly its social meanings and, more importantly, is state constraint part of these meanings? Is the direct interpretivist argument, with all the evidence provided there, accurate? According to the sceptic (see also 2.3), without any fixed, external criteria, the interpretivist cannot offer a definitive or correct picture of his own culture and society. He exposes shared understandings and self-perceptions and backs his position

with certain facts about the historical development of his society. Nevertheless, it is impossible to determine what is a prevalent conception and what is a historical aberration or transient phenomenon, without any fixed external criteria. One could, for example, argue that the distinction between the political and the non-political does not reflect the true character of democratic societies and that Rawls' presentation of full political participation as one among many conceptions of the good is based on a gross misunderstanding of the political process.

For the non-sceptical interpretivist it is not impossible to distinguish the good (true) interpretation from the bad one on the basis of the evidence presented. Nevertheless, if he wants to speak the language of persuasion and appeal even to the sceptics in his audience, he is likely to find that the direct interpretivist argument is not always the most effective means of exerting influence. For this reason, Walzer suggests, as we saw in Chapter 2, that the interpretivist argument must have an edge, must be radical enough to capture the imagination of the audience. Rawls has tried another way of making his interpretivist prescriptions attractive, one which is less likely to compromise the (interpretivist) truth (however uninteresting or mundane it may appear) in order to leave a lasting impression on the audience. This new interpretivist route is the constructivist argument which I will now examine closely.

## 5.2 The Constructivist State Constraint Argument

The first level of interpretation in the constructivist argument involves the design of two "model conceptions" which incorporate ideas latent in the thick language of democratic societies and familiar to their members. Having established these conceptions as the terms in which we could understand the problem of social justice, Rawls constructs a hypothetical situation, the Original Position, which incorporates certain elements of the democratic process.

As we shall see, the Original Position is, to a large extent, simply a matter of presenting Rawlsian prescriptions as the outcome of a familiar political procedure: delegates convene to decide on principles which will be binding on themselves and on their constituents, principles which will govern the institutional basis of their society. This hypothetical situation is meant to resemble a constitutional convention or a legislative assembly so that it can appeal to our moral/political intuitions and thus gain the same degree of authority and respectability which we presumably recognise in such law-making procedures. I believe that this is why Rawls calls the Original Position a "device of representation".<sup>18</sup>

At the same time, however, the Original Position can be seen as itself a device of interpretation, as a precursor to the constitutional, legislative and judicial levels of social interpretation. It binds together the model conceptions in a manner appropriate to Western democratic societies and what it, therefore, yields is not simply principles of justice but

a more coherent procedure of further specification of these principles. In 6.2 I will argue that Rawls' intention to appeal to our intuitions by way of presenting familiar pictures drawn from public life in democratic societies, does little to demonstrate that his prescribed principles are really implicit in the moral and political practices of these societies.

The starting point of this discussion will have to be the first one of Rawls' model conceptions, namely the idea of the "moral person" or of the "person qua citizen". Rawls makes it clear from the outset that this idea is not founded on a metaphysical doctrine, but is "a political conception of the person" latent in Western political culture.<sup>19</sup>

In his recent writings Rawls offers a portrayal of the person qua citizen drawing from characteristically Western self-perceptions, moral intuitions and shared meanings. His first point is that citizens "...are free in that they conceive of themselves and of one another as having the moral power to have a conception of the good."<sup>20</sup> Furthermore, the citizens' public identity as free persons is not affected by changes over time in their conception of the good and, therefore, there can be no loss "...of what we may call their public identity as a matter of law."<sup>21</sup> Rawls hastens to add that this is not the case in societies where there is a different political conception of the person. As he puts it, "[w]e can imagine a society (indeed, history offers numerous examples) in which basic rights and recognized claims depend on

religious affiliation, social class and so on."<sup>22</sup>

A second characteristic of persons qua citizens is that they have the "capacity for an effective sense of justice."<sup>23</sup> This means that they can accept, live by and defend principles governing the basic institutional structure of society. The capacity for a sense of justice is not connected with some essentialist human urge to justify one's actions, ends and aspirations to others, but with two self-perceptions which are widely shared in Western democratic societies.

First, as Rawls himself puts it, "...citizens view themselves as free in that they regard themselves as self-originating sources of valid claims." Rawls makes it clear that this is not the case in societies with different political conceptions, where individuals' claims have weight only insofar as they can be "derived from peoples' duties and obligations owed to society, or from their ascribed roles in the social hierarchy justified by religious or aristocratic values."<sup>24</sup>

Second, citizens are regarded as free in the sense that they are capable of taking responsibility for their ends. The idea here is that, given certain conditions of equality (which the interpretivist seeks to determine and incorporate into his principles of justice), individuals "...are thought to be capable of adjusting their aims and aspirations in the light of what they can reasonably expect to provide for. Moreover they are regarded as capable of restricting their claims in matters of justice to the kinds of things the principles of justice allow."<sup>25</sup> So responsibility here means bearing the

cost of pursuing one's aims under certain conditions of equality, and accepting these conditions as mutually binding. Again this can be regarded as a point of contrast with existing societies where the claims of the members of specific groups or social classes are not assessed in this manner. In an aristocracy, for instance, the members of the ruling class are spared at least some of the cost of pursuing their aims and aspirations while members of lower classes can only pursue specific aims even if they can afford the cost of higher (more expensive) ones.<sup>6</sup>

In connection with this model conception Rawls discusses the issue of primary goods, that is, the social conditions and all-purpose means which enable individuals to pursue their determinate conceptions of the good and to develop and exercise their two moral powers. To determine these goods (which are to be distributed, directly or indirectly, through the institutional structure) "...we must look to social requirements and to normal circumstances of human life in democratic society."<sup>26</sup>

As in the direct argument, the determination of primary goods invokes knowledge of the general circumstances and requirements of social life in Western societies. But in the constructivist argument it does so only in the light of "[the political] conception of the person". This means that instead of getting the index of primary goods only in the Walzerian

---

<sup>6</sup>Rawls' idea of contrasting the self-perceptions that are prevalent in Western societies with those that are predominant elsewhere does not sit very well with his interpretivism since it suggests indirectly that these are categorical differences between essentially distinct, immutable worlds.

manner (by looking for the shared meaning of social goods in language, literature and the Western history of institutional development), Rawls crosschecks the findings of a Walzerian investigation with the features of his first model conception. It seems that, for Rawls, the latter method renders his argument more coherent and defensible.

Nevertheless, one might argue that deriving the meaning of social goods from a conception of the person (even from a specifically political, parochial one) is not philosophizing as an interpretivist. This, the critic might say, is foundationalist deductionism in disguise, a way of deriving everything from a specific view of human nature which is presented as an interpretivist finding.

I believe that the best response to this criticism that is available to Rawls is the following: The design of the model conceptions, and of the conception of the "moral person" in particular, is itself an interpretivist project carried out in parallel with the Walzerian examination of the meaning of social goods. The two findings are subsequently juxtaposed in order to determine whether they are congruent with each other. The underlying idea is that by having a political conception of the person the interpretivist can crosscheck the index of social goods which he has uncovered and achieve a coherence of the whole of shared meanings which he examines. To put it differently, it is the design of a conception of the person qua citizen (based on shared self-conceptions and familiar moral/political ideas like freedom and responsibility) that confirms the findings of the Walzerian (i.e. the direct)

method.

To clarify this point it is useful to compare interpretivist political philosophy with Schleiermacher's theory of literary hermeneutics. According to the latter, the literary critic must reconcile all the different parts of a text and show how these work together to compose a well-integrated meaning. If certain parts of the text contradict others, the initial presumption is that one or some of the parts have been misunderstood. It is the idea of the comprehensive unity of meaning that guides interpretivist's critical understanding and gives him a sense of direction.<sup>h</sup> The idea of the person qua citizen purports to do exactly that; to enable the political philosopher to choose the "proper" interpretation of the meaning of social goods.

But even if this method is not incongruous with interpretivist meta-theory, it simply cannot serve as the missing criterion by which one can arrive at a single, accurate interpretation. In Rawls' case, it simply transfers the problem: Since the model conception is itself the outcome of an interpretivist project its accuracy is now in doubt. Ultimately the whole argument rests on the optimistic assumption that an interpretation of the character of the person qua citizen in Western societies is rather

---

<sup>h</sup>As Georgia Warnke points out in her Justice and Interpretation (Polity Press: Cambridge, 1992), p. 21, even deconstructionists implicitly concede that interpretation involves a holistic understanding of the text. As she puts it, "[d]econstructionists must have an idea of what the intention or meaning is that is undermined by the text's language and structure and it must have acquired this idea from an appreciation of how various parts of the text at issue are meant or attempt to compose a unified whole."

straightforward.

Rawls' second model conception is the idea of a "well-ordered society". Rawls calls this "the overarching, fundamental intuitive idea" and claims that it encompasses certain ideas about the ideal social arrangement which are implicit in democratic political culture. These ideas are presented as features of the second model conception.

First, there is "cooperation" among the members of society. Cooperation is regarded as distinct from "merely socially coordinated activity" and is guided by publicly recognized rules and procedures regulating the conduct of those who are cooperating.<sup>27</sup> Second, cooperation involves the idea of "fair terms of cooperation", that is, terms that each participant may accept provided that everyone else does the same and which specify "an idea of reciprocity or mutuality: all who are engaged in cooperation and who do their part as the rules and procedures require are to benefit in some appropriate way as assessed by a suitable benchmark of comparison."<sup>28</sup> The fair terms of cooperation are turned into principles of justice when they are predicated of the basic institutional structure of society.

An implication of the existence of fair terms of cooperation is that the conception of justice is publicly known and generally acknowledged. It follows from Rawls' definition of cooperation that if this condition does not hold, cooperation itself becomes problematic if not impossible. This point will be discussed at length later on.

Finally, the idea of social cooperation "...requires an idea of each participant's rational advantage, or good."<sup>29</sup> Each participant tries to achieve some conception of the good through social cooperation and therefore the significance of such conceptions to the well-ordered society is great. This point as well as the previous ones are presented in this rather sketchy manner not because a more detailed picture would be difficult to compose, but because Rawls' plan involves breaking down the interpretive project into smaller tasks whose results are then combined so that they interact with one another to produce the intended effect. More about the rationale behind constructivist methodology will be said along this presentation of the constructivist argument.

The link between the two model conceptions presented above is the hypothetical process Rawls calls the Original Position: An unspecified number of "moral persons" gather together to decide on principles for a well-ordered society. These fictitious characters are thought to be the representatives of actual individuals, that is, of individuals with a set identity (gender, colour), with particular interests and conceptions of the good.

The Original Position incorporates two ideas which Rawls finds characteristic of Western democratic societies: First, representation (i.e. having delegates, whose purpose is to serve the interests of their constituents, confer in order to reach common decisions binding to all) and, second, proceduralism (the idea that whatever principles are chosen or decisions reached in a certain kind of situation are proper or

"just").<sup>i</sup>

It is, I think, clear that the Original Position is itself meant to be a distinct interpretive element, a third model conception which could have been dubbed "proper procedure". The Original Position is the proper procedure because it models the ideas of representation and proceduralism which characterize the legislative and judicial procedures in modern democratic societies.<sup>30</sup>

One of the most widely advertised features of the Original Position is the idea of the "veil of ignorance". According to this idea, the delegates who take part in the proceedings in the Original Position are not to know the particular identity (the gender, colour and conceptions of the good) of those whom they represent. This point raises the following question: In what sense are the delegates to mind the interests of their constituents if they know nothing about them? The answer here is this: Even behind the veil of ignorance the delegates have some information about their constituents. They know that like themselves they have two moral powers, the power to have a conception of the good and the power to have an effective sense of justice. Their mission, therefore, is to protect these powers, to treat them as three distinct "interests": a higher-order interest in forming, revising and rationally pursuing a conception of the good, a higher-order interest in having an effective sense of justice, and a high-order interest "...in protecting and

---

<sup>i</sup>Of course, for the interpretivist "just" belongs to the thick moral/political vocabulary of the society in question.

advancing their conception of the good as best they can, whatever it may be."<sup>31</sup>

The question that arises at this point is why Rawls imposes the veil of ignorance on the moral persons who take part in the Original Position instead of asking them to decide on principles according to their own "interests" (which coincide with those of actual people seen through the veil). There are, I believe, three reasons for this.

First, if Rawls were to derive principles of justice from the conception of the moral person alone, he would be accused of contradicting his proclaimed interpretivism. Although the first model conception is the result of an interpretivist investigation, deducing principles from this conception alone would be using it as a foundation.<sup>j</sup> The design of such a character may be useful to the interpretivist philosopher, but it cannot encapsulate all the predominant shared meanings, moral ideas and principles that are latent in a living social and political tradition. In other words, no single model conception, and, more specifically, none of those proffered by Rawls, can lift the weight of a whole interpretivist theory.

---

<sup>j</sup>Of course, according to the taxonomy of Chapter 2, the foundations are universal essences and, therefore, Rawls' model conceptions (being self-perceptions and ideas that are dominant in a particular society at a given time) can never be regarded as such. Nevertheless, the idea of deducing all moral/political principles from a single conception has a distinct foundationalist flavour. This does not mean that if it were applied by Rawls, it would make his position foundationalist in character. Rather, the point is that it would make his interpretivist argument very superficial and would raise doubts about his commitment to interpretivist meta-theory--the latter being some source for concern for Rawls, who has repeatedly denounced foundationalist philosophy.

Rawls is conscious of this point. The constructivist approach enables him to let the model conceptions (which are separate, not very detailed interpretive pictures) interact with each other and thus avoid the controversy that a single, detailed interpretive picture would give rise to.

Second, for Rawls, the veil of ignorance itself is the modelling of a certain moral outlook latent in the political practices of Western societies which roughly holds that the distribution of certain (social) goods among individuals or groups of individuals should not reflect the relative advantage (physical, social, economic or arithmetic) of any of these individuals or groups. There clearly is a certain affinity between this idea and two others which have been built into the model conceptions: The idea of "responsibility for one's aims" and the idea of "fair cooperation". The three of them are closely related and complement each other to form the crux of the Rawlsian conception of justice.

Third, without the veil of ignorance, the ideas of representation and proceduralism would have to be left out of the Original Position. For the interpretivist Rawls this would be unthinkable. In his view, representation and proceduralism are two very prominent features of the political conception of Western democracies and cannot possibly be omitted from his interpretivist picture. He uses them to construct a hypothetical situation which facilitates the interaction between the two model conceptions in a manner that resembles the three actual levels of social interpretation and conflict-resolution in Western societies, namely the constitutional,

the legislative and the judicial.

Nevertheless, one might argue that there is only a formal resemblance between the Original Position and the actual constitutional and legislative procedures. Representation, for example, does not help to shape the actual outcome of the hypothetical choice situation and would not be included were not Rawls keen on incorporating into the Original Position some of the outward characteristics of the democratic process. In my view, this is not the case.

The ideas underlying representation (e.g. equality, reciprocity, responsibility) in parliamentary democracies are already built into the two model conceptions and contribute to the formulation of the principles of justice. It is, however, true that formal resemblances to actual processes are important to Rawls for reasons that are not related to the pursuit of philosophical truth: As somebody who speaks the language of persuasion, Rawls is interested in presenting the prescribed set of principles of justice as the outcome of a (hypothetical) process whose features are familiar to his audience. To reiterate a point made earlier, the constructivist argument is also (perhaps primarily) a political argument, an attempt on Rawls' part to make his, already formed, interpretivist conclusions attractive to more people.<sup>k</sup> This must have been one of the considerations that

---

<sup>k</sup>It is important to note that an interpretation that captures the prominent moral/political features of a particular society is very likely to appeal to most of its members. However, this does not mean that the interpretivist offers the interpretation that is more likely to satisfy the widest possible majority within society. An interpretivist position is not informed by public opinion polls nor is it

shaped the Original Position of the later Rawls.

Let us now turn to the formulation of the state constraint principle by the participants in the Original Position. Each participant must distinguish between three kinds of considerations: First, considerations concerning the particular conception of the good which his constituents hold and which is unknown to him personally, second, considerations concerning his constituents' capacity for choosing, revising and pursuing a conception of the good, and third, considerations concerning their capacity for a sense of justice.

At this point a list of liberties drawn up by the interpretivist (who has already searched for ideas latent in the practices of Western democracies and for the presumed meaning of public goods) is made available to the participants in the Original Position. The delegates are supposed to crosscheck the findings of the interpretivist philosopher by justifying each one of the liberties in the list in terms of any of the three kinds of considerations mentioned above. In his "The Basic Liberties and Their Priority" Rawls demonstrates how the participants in the Original Position examine and accept liberty of conscience. First of all, they decide in favour of liberty of conscience because they do not know whether the conception of the good of their constituents is a minority or a majority conception, that is, whether the

---

supposed to predict the outcome of a process of negotiation between competing interest groups or moral outlooks. This point is discussed in more depth later on.

supporters of that conception are more numerous and therefore more capable of taking coercive action against heterodox minorities or less numerous and therefore more vulnerable. The delegates endorse liberty of conscience because they think it minimises the risk of their constituents becoming the victims of moral coercion.

Second, and in connection with the capacity for a conception of the good, the delegates take the view that liberty of conscience enables their constituents to revise their chosen conception. As Rawls puts it, "[t]here is no guarantee that all aspects of our present way of life are the most rational for us and not in need of at least minor if not major revision."<sup>32</sup> In other words, re-evaluation and revision of chosen conceptions of the good is an aspect of the first moral power. It is for this reason that the delegates in the Original Position deem liberty of conscience suitable for their constituents. Furthermore, they find that their constituents may want to strive to appreciate **why** their conceptions of the good are true or valuable or suitable for them. They may want, as Mill would say, to make their conception of the good "their own" instead of accepting it ready-made from their close environment or social peers.<sup>33</sup> Liberty of conscience enables them to achieve this.

Now, liberty of conscience cannot be justified in terms of the capacity for a sense of justice, but this is not necessary for it to be incorporated into the scheme of equal liberties. As has been mentioned, the delegates endorse the liberties which promote or secure the exercise of any one of

the two moral powers (capacities). The equal political liberties and freedom of thought, for instance, are not directly related with the capacity for a conception of the good, but with the capacity for a sense of justice. As Rawls puts it, they "...secure the free and informed application of the principles of justice, by means of the full and effective exercise of the citizens' sense of justice, to the basic [institutional] structure".<sup>34</sup>

The participants in the Original Position also seek to establish that the set of principles of justice they decided on is as a whole advantageous to all and, therefore, acceptable to their constituents. Rawls lists three reasons why the principles of justice (and not simply a particular liberty or set of liberties) are advantageous to the constituents of the participants in the Original Position. First, they enhance self-respect since they entail a situation in which citizens need not feel fear or shame for adhering to a particular conception of the good. Second, they contribute to social stability by providing ample space for the pursuit and achievement of many different conceptions of the good and, finally, they conduce to the development of an all-inclusive, comprehensive good, "a social union of social unions".<sup>35</sup>

It should be noted here that the idea of the principles of justice being accepted by all on the basis of being advantageous to all is not fully congruent with interpretivist meta-theory. The interpretivist's objective is not to accommodate the interests and loyalties of all members of society but to set out moral/political principles which he

finds dominant in the society he investigates. Whether or not his prescriptions will be accepted by the majority of citizens as advantageous is not something that he takes into consideration during his investigation. Rawls is not totally oblivious of this point. According to his argument, the constituents cannot reject or ignore the sense of justice that took shape in the Original Position because they are "reasonable", that is, they do not only know and pursue what is good for them personally, but also will to conform to principles of justice.<sup>1</sup>

What comes out of the interpretivist apparatus of the Original Position is the same liberty and harm principles presented during the discussion of the direct argument. What is, however, different is the argument in favour of the priority of the liberty principle over the second principle of Rawlsian justice. As the commentator Rex Martin notes, the basic liberties are now regarded as necessary means or essential parts of realizing the two higher-order interests whereas the rest of the primary goods (opportunities, powers, income and wealth) are associated with lesser interests and appear in a quite different argument (the idea of a collective asset) that makes no clear reference to the two moral powers of the person qua citizen. Liberty can thus be thought to have a sort of generic priority over the other primary goods.<sup>36</sup> This priority is not latent or implicit in moral and political practices, but is a result of the interaction between the

---

<sup>1</sup>As we shall see later on, even this disinterested acceptance of the prescribed principles is no proof that the philosopher's interpretation is accurate.

different interpretive elements of the constructivist approach.

In order to understand the process of specification of the liberty and harm principles, it is important to further clarify the interpretive role of the Original Position. As has been mentioned, the Original Position is supposed to bare the outward characteristics of a legislative assembly because through it the interpretivist does not only aspire to articulate regulative principles implicit in Western moral/political practices, but also, by doing that, to align the different levels of law-making (i.e. the constitutional and the legislative), to have the outcome of the first level constrain the participants in the next level. To put it differently, Rawls puts together his interpretive findings to construct the initial, hypothetical level of law-making that precedes and constrains all lower levels of law-making and thus renders the actual process of law-making in democratic societies more coherent and congruent with their social meanings. Parties in the Original Position are constrained by the veil of ignorance or rather by what Rawls calls "the reasonable conditions built-in the Original Position" in order to decide on principles of justice which will then constrain the parties taking part in a constitutional convention. The latter yields a constitution which along with the principles of justice constrains the parties in the legislative stage. Both of the actual stages of law-making are levels of specification of the principles of justice and, like the

Original Position, they are different levels of social interpretation.

In attempting to specify the liberty and harm principles formulated in the Original Position one must, therefore, answer the following question: "What more particular liberties, or rules of law, are essential to secure the free, full and informed exercise of the moral power for a sense of justice ?"<sup>37</sup> As Rawls asserts, one must not proceed "...from a general definition that singles out these liberties, but from what the history of constitutional doctrine [itself a tradition of social interpretation] shows to be the fixed points within the central range..." of each particular liberty.<sup>38</sup>

In "The Basic Liberties and Their Priority" Rawls examines the cases of freedom of speech and freedom of the press, both of which have been traditionally regarded as requirements of the freedom of thought and the equal political liberties which themselves secure, according to the constructivist argument, the exercise of the moral power for a sense of justice. The history of constitutional doctrine shows that there are three fixed points within the central range of the freedoms of speech and the press. First, there is no such thing as the crime of seditious libel, second, there are no prior constraints on freedom of the press and, third, there is full protection of revolutionary, subversive doctrines.<sup>39</sup>

In connection with the issue of seditious libel Rawls states that

[t]he history of the use by governments of the crime of seditious libel to suppress criticism and dissent and to maintain their power demonstrates the great significance of this particular liberty [i.e. freedom of speech] to any adequate scheme of basic liberties. So long as the crime exists, the public press and free discussion cannot play their role in informing the electorate....Thus the great importance of N.Y. Times vs. Sullivan in which the Supreme Court not only rejected the crime of seditious libel but declared the Sedition Act of 1798 unconstitutional....It [i.e. the Sedition Act] has been tried, so to speak, by the court of history and found wanting.<sup>40</sup>

It is obvious that Rawls finds the use of the crime of seditious libel by Western governments as a historical aberration rather than a prominent feature of the public morality of Western democracies. What is important to note here is that Rawls accepts the ruling of the U.S. Supreme Court both as a valid, direct interpretivist argument and as the outcome of a procedure in which the participants are constrained by the principles of justice. Once again, the direct argument is incorporated into the constructivist position and is constrained by it.<sup>m</sup>

Until this point Rawls would have us believe that what constituted harm in a democratic society was the violation of the basic and secondary personal liberties as well as the rights of the individual connected with the second principle

---

<sup>m</sup>It should, however, be noted that Rawls' view on the role of the U.S. Supreme Court is not very clear. Supreme Court judges are not constrained in the same way that participants in a legislative assembly or a constitutional convention are. Their position is closer to that of the interpretivist philosopher himself, who interprets without any constraints imposed on him, than to that of the congressman or member of parliament.

of social justice. What he now avers is that liberties are also self-limiting. In connection with freedom of speech, for instance, he states that "...there must be some point at which political speech becomes so closely connected with the use of force that it may be properly restricted..."<sup>41</sup> In other words, there is a point beyond which political speech constitutes harm, not to particular individuals (since no-one's individual rights and liberties are necessarily violated) but implicitly to every citizen since the stability and existence of the (just) institutional arrangement is under threat. Rawls avers that a situation in which the existence of just institutions is threatened by political speech, and especially by the advocacy of revolutionary or subversive doctrines, is hardly imaginable. After all, if political speech is guaranteed, the serious grievances of any social groups do not go unrecognized or suddenly become highly dangerous. Nevertheless, as Rawls himself notes, there is a certain rule, known as the "clear and present danger rule", which has an important place in the history of constitutional doctrine in the U.S., and which concerns the content of political speech. According to one formulation,

The question...is whether the words are used in such circumstances and are of such nature as to create a clear and present danger that they will bring about the substantive evils that Congress has a right to prevent. It is a question of proximity and degree.<sup>42</sup>

The use of certain words in political speech and under

specific circumstances may, therefore, be regarded as harmful in the sense that it incites others to take action which will destroy just institutions and deprive everyone of their basic liberties. But if the citizens of a democratic society, whose basic structure is governed by the principles of justice, can be goaded into, or rather, talked into taking action to subvert just institutions, if, in other words, consensus on the Rawlsian conception of justice crumbles or the principles of justice are rejected by public opinion and given that, according to Rawls, the interpretivist must prescribe principles which are acceptable to a large majority of citizens, it follows that any substantial shift in public opinion suggests the need for revisions and reforms rather than the enforcement of an outdated conception of justice and the censoring of any expression of strong dissent.

The reason why Rawls does not arrive at this conclusion, is that he does not regard sudden changes in public opinion as an indication (or as sufficient indication) that social meanings, the widely shared, tacitly acknowledged ideas and principles, have changed. What, therefore, warrants censorship in particular circumstances is the need to protect institutions governed by principles which (still) capture these prevalent meanings.<sup>n</sup> This point is fully compatible with

---

<sup>n</sup>There are two more interpretations of Rawls' argument in favour of censoring certain uses of political speech. According to the first, Rawls is saying that the dissemination of subversive ideas should be restricted because it undermines social peace. Commentators who attribute this argument to Rawls find a strong Hobbesian element in his more recent work. For instance, Kukathas and Pettit claim Rawls' position has assumed a "decidedly Hobbesian flavour, since he now ties his conception of justice...[to] order." [in Rawls: A Theory of

the distinction (made in Chapter 2) between the philosophical and the political. Political discontent like political apathy (see 5.1) is no sure indication that social meanings have changed, that the prescribed principles are no longer implicit in practice or tacitly accepted. In fact, it is often the case that rampant rioting and outbreaks of civil disobedience are caused by what are perceived by the protestors as deviations, on the part of state officials, from the prevalent, already institutionalized conception of justice.

But what exactly are the circumstances under which the restriction of certain uses of political speech is warranted? The problem here is that if the answer to this question is not clear enough, the state constraint principle loses its significance and becomes open to various interpretations. What can now constitute harm is not only the violation of the rights and liberties of others, but conduct (political speech) that is deemed likely to lead to the collapse of the

---

Justice and its Critics (Polity Press: Cambridge, 1989), p. 140] In my view this argument is not sufficiently well incorporated into Rawls' interpretivist position and, if it is to be attributed to Rawls at all, should be seen as part of his political (as opposed to strictly philosophical) argument.

According to the second interpretation, Rawls is saying that the principles of justice are essential characteristics of Western democratic tradition. Any serious threat to the institutions governed by these principles threatens the democratic tradition itself. Such arguments are regularly employed by those who speak the language of persuasion. (For instance the self-proclaimed defenders of the "American way" who brand their political rivals' positions as "unamerican" in order to discredit them, use this type of argument.) In terms of the taxonomy of Chapter 2, such arguments make for bad explanationist philosophy since they have a strong normative flavour. If Rawls' argument in favour of qualified censorship is indeed such an argument, he employs it in order to make his overall (interpretivist) position more attractive to his audience.

institutions of justice. Rawls hastens to add that freedom of speech can only be restricted in the name of liberty. This point hinges exclusively on the assumption that the subversion of just institutions will ultimately deprive most citizens of their liberty. Although this may be a valid assumption, there have to be safeguards against the suppression of dissenting minorities. In order to stress the non-partisan character of his state constraint position Rawls comes up with one simple principle according to which

...free political speech cannot be restricted unless it can be reasonably argued from the specific nature of the present situation that there exists a **constitutional** crisis in which democratic institutions cannot work effectively...<sup>43</sup>

According to Rawls, a constitutional crisis is a situation in which the institutional structure has broken down and where the distribution of political (primary) goods becomes problematic. In those societies with a long democratic tradition such situations can arise when their "...people and institutions are simply overwhelmed from the outside."<sup>44</sup> Any restrictions on liberty are, therefore, meant to make possible the transition to the original state of effective operation of the just institutions when outside pressure is relaxed. Rawls concludes his discussion of the clear and present danger rule stating that "for practical purposes, in a well-ordered democratic society under reasonably favourable conditions, the free public use of our reason in questions of political and social justice would seem to be absolute."<sup>45</sup> This conclusion

is presumably reached through a direct argument showing that the near-absolute character of the constraint principle is part of the conception of justice which is dominant in modern democracies.

Before I conclude this chapter, I should touch on the interpretivist Rawls' distinction between what he calls "an overlapping consensus" and "a modus vivendi", a distinction which is meant to clarify his interpretivist position but which itself needs further clarification.

In his recent work, Rawls has claimed that the philosopher must propose a conception of justice which best captures the social meanings of modern, democratic societies and at the same time a conception on which the vast majority of citizens can agree. Agreement on interpretivist principles that could govern the institutional structure of society is what Rawls calls an "overlapping consensus", a "consensus including the opposing religious, philosophical and moral doctrines likely to thrive over generations in a society effectively regulated by that conception of justice."<sup>46</sup> For Rawls, the important point here is that while his conception of justice is not identified with any particular conception of the good (or comprehensive doctrine) those supporting different conceptions each have different grounds for accepting it. Rawls contrasts the idea of an overlapping consensus with what he calls a modus vivendi, that is, with the idea of justice as a compromise between comprehensive doctrines or as an arrangement justified in purely

instrumental terms.

The first question that arises here is whether Rawls understands modern, Western societies (and the U.S. in particular) as examples of modus vivendi arrangements or as societies where there exists an agreement on political principles without an agreement on the grounds for them. On many occasions Rawls talks of an overlapping consensus as something that is yet to be achieved. As he puts it, "...we hope to make it possible for all to accept the political conception as true, or as reasonable, from the standpoint of their own comprehensive view whatever it may be."<sup>47</sup> But elsewhere Rawls talks of past and present constitutional stability (at least in the U.S.) and avers that "...history shows that democratic institutions are quite resilient in the U.S.". He also refers to historically "successful democratic constitutions" and to "constitutions that work".<sup>48</sup> Are we to infer from this that an overlapping consensus is already a reality or that constitutional stability and a modus vivendi are not mutually exclusive situations? Clearly if he thought the overlapping consensus a reality, Rawls would not be seeking to achieve it. This, however, does not necessarily mean that he understands Western democracy as a modus vivendi, or, to use his term, a "mere modus vivendi" (what he also refers to as "unwilling, resentful compliance"). In a fast changing, increasingly pluralistic society Rawls sees the need for a re-interpretation, a fresh articulation of the moral/political principles implicit in moral/political practices. It is clear, however, that in his view these

principles must be articulated and presented in such a way as to make them acceptable to the largest possible majority of citizens. Citizens must be enabled to find reasons within their own moral outlook to accept the proposed political conception of justice. Given that justice is a specifically **political** conception and that social divisions are not (yet) too deep, Rawls hopes that the supporters of comprehensive doctrines will not have reasons to reject it. As for those who do not belong to any organized group of adherents of a particular doctrine (a church, a club, a cultural or ethnic community etc.) Rawls seems to believe that they can be regarded as supporters of neo-Kantian foundationalist philosophy (in which case they would accept Rawlsian interpretivist justice because it prescribes principles very similar to those of neo-Kantian justice).

It should be emphasized that Rawls seeks to make his interpretivist position acceptable by presenting it in a particular way and not by altering the content of his regulative principles so that he satisfies the supporters of as many conceptions of the good as possible. It is true that on many occasions he seems to be suggesting that the most accurate interpretation is the one which achieves the highest level of acceptability, or, alternatively, the one which locates those elements that are common to most of the comprehensive doctrines held by members of democratic societies. And yet Rawls does not manipulate his interpretivist position in order to make his audience agreeable to it (although he does present it in a purportedly

persuasive manner) and does not think of it as the area where most disparate ideals of the good life overlap because if that were the case he would not be talking of a specifically political conception.<sup>o</sup>

One could, however, argue, as I believe Rawls does, that agreement on general principles governing the institutional structure of society is itself a requirement of the Western conception of justice. If this is so, how can the interpretivist philosopher prescribe principles without heeding the opinions and expressed views of the members of the society he investigates? How can he ignore their interests and loyalties?

The answer has already been given: the social meanings and self-perceptions which the interpretivist takes into consideration in the formulation of regulative principles cannot be found at the level of public opinion. Neither the interests nor the opinions of actual individuals can affect social meanings or the ideas implicit in practice. Mechanisms through which consensus on moral/political principles are to be achieved can be prescribed by the interpretivist, but cannot be built into the interpretivist argument itself and cannot influence the outcome of the interpretive project. As

---

<sup>o</sup>The use of the term "overlapping consensus" is, I believe, unfortunate because it obscures one of Rawls' main points: justice is a specifically political conception independent of all comprehensive doctrines and not an area where comprehensive doctrines overlap. However, I believe that Rawls consciously adopts this term in order to make a favourable impression on the members of minority (mainly cultural and religious) groups who resentfully conform to the political conception he has identified. Judging from the popularity that the term now enjoys, this political move has been successful.

I have already mentioned, the very construction of the Original Position confirms this point. The opinions and particular interests of the constituents are not taken into consideration in the formulation of the principles by their representatives. But then, the question arises, why is Rawls so keen on presenting the principles of justice as advantageous to all, as principles that everyone could find reasons to accept? Rawls is not only interested in disclosing the truth but in exerting influence, in persuading his audience and outmanoeuvring his critics. He, in other words is a politician who has proven to be at least as competent as he is a political philosopher.

The constructivist approach of building interpretivist findings into a hypothetical situation resembling a legislative assembly where principles of justice are decided is a political argument in so far as it makes use of imagery and terminology which is familiar to the citizens of Western democracies in an attempt to make the Rawlsian position attractive to them. But it is also a philosophical argument. Instead of offering a single, direct interpretivist argument in support of the principles of justice, Rawls combines the findings of separate interpretivist projects (e.g. the notion of the moral person, the idea of a well-ordered society, the notions of representation and proceduralism) into a construction which yields principles of justice. In other words, he divides his interpretivist project into smaller ones whose results he then combines to reach his final conclusion. Of course, this approach cannot provide a criterion of

accuracy (i.e. a criterion by which to distinguish true social meanings from mere accidents or from tendencies which have already expired) since it is dependent on direct interpretivist arguments. One could, therefore, conclude that from a meta-theoretical point of view neither of the two interpretivist approaches is preferable to the other.<sup>p</sup>

Following the discussion and analysis of foundationalist and interpretivist state constraint, I should now be turning to explanationist state constraint. There is, however, an important reason why I should do so after I discuss the neutrality debate: since the debate itself is an argument between normativists, explanationist state constraint (which, unlike the two views already examined, is a non-normative position) is not involved in it and can, therefore, be examined separately.

In order to take a close look at the neutrality debate it is important to re-insert "neutrality" in the place of "state constraint" for the simple reason that this is the term most of the participants use. Recognizing the difficulties caused by the contextuality of "neutrality" (see Chapter 1) some of them give elaborate definitions of state neutrality in an attempt to avoid misunderstandings.

---

<sup>p</sup>It could be argued, however, that the way the particular interpretivist findings are incorporated into the Original Position is arbitrary because it is devised by the philosopher himself. This would be a reason to prefer the direct approach to the constructivist. Rawls' response to this would be that his interpretivist findings interact in a way that is itself implicit in the practices of Western democratic societies.

## Notes

1. John Rawls, "Kantian Constructivism in Moral Theory", Journal of Philosophy, 76 (1980), p. 512.
2. John Rawls, "Justice as Fairness: Political Not Metaphysical", Philosophy and Public Affairs, 14 (1985), p. 228.
3. For instance C. Kukathas and P. Pettit call the recent developments in Rawlsian philosophy "The Self-Critique" in Rawls: A Theory of Justice and its Critics (Polity Press: Cambridge, 1989), pp. 119-151.
4. John Rawls, "The Idea of an Overlapping Consensus", Oxford Journal of Legal Studies, 7 (1987), p. 13.
5. Ibid., pp. 26-27.
6. John Rawls, "The Basic Liberties and their Priority", in S. McMurrin (ed.), Liberty, Equality and Law (University of Utah Press: Cambridge, 1987), p. 5.
7. Ibid., p. 58.
8. John Rawls, supra note 4, p. 5.
9. Ibid., p. 14.
10. John Rawls, supra note 6, p. 28.
11. John Rawls, A Theory of Justice (Oxford University Press: Oxford, 1989), p. 330.
12. John Rawls, supra note 4, p. 9.
13. John Rawls, "The Priority of Right and Ideas of the Good", Philosophy and Public Affairs, 17 (1988), p. 252.
14. John Rawls, supra note 6, p. 45.
15. John Rawls, supra note 4, p. 14.
16. Ibid., p. 8.
17. John Rawls, supra note 6, p. 37.
18. John Rawls, supra note 2, p. 229.
19. Ibid., p. 243.
20. Ibid., p. 240.
21. Ibid., p. 241.

22. Ibid., p. 241.
23. John Rawls, supra note 1, p. 525.
24. John Rawls, supra note 2, p. 243.
25. Ibid., p. 244.
26. John Rawls, supra note 6, p. 15.
27. John Rawls, supra note 2, p. 232.
28. Ibid., p. 232.
29. Ibid., p. 232.
30. John Rawls, supra note 1, p. 525.
31. Ibid., p. 525.
32. John Rawls, supra note 6, p. 27.
33. Ibid., p. 28.
34. Ibid., p. 30.
35. Ibid., p. 34.
36. Rex Martin, Rawls and Rights (University Press of Kansas: Lawrence, 1985), p. 108.
37. John Rawls, supra note 6, p. 57.
38. Ibid., p. 57.
39. Ibid., p. 57.
40. Ibid., p. 57.
41. Ibid., p. 58.
42. Ibid., p. 63.
43. Ibid., p. 64.
44. Ibid., p. 70.
45. Ibid., p. 71.
46. Ibid., p. 71.
47. John Rawls, supra note 4, p. 13.
48. John Rawls, supra note 6, p. 69.

## **Part III: The Neutrality Debate**

## CHAPTER 6

### State Neutrality

The state constraint positions discussed in the last three chapters are usually placed under the single heading of "state neutrality" without distinction of meta-theoretical background. The political philosophers I presented as authors or defenders of these positions are self-proclaimed neutralists. David Richards, for instance, thinks of his position as a "moral neutrality with regard to the many disparate visions of the good life". According to him, neutrality holds that "the concern embodied in the idea of human rights is not with maximizing the agent's pursuit of any particular lower-order ends, but rather with respecting the higher-order capacity of the agent to exercise rational autonomy in choosing and revising his ends, whatever they are."<sup>1</sup> In other words, the moral principles to be predicated of the institutional structure (i.e. of the state) are **neutral** between particular conceptions of the good. In similar fashion Ronald Dworkin states that "the government must be **neutral** on what might be called the question of the good life."<sup>2</sup> Will Kymlicka defends what he calls "liberal neutrality" against certain criticisms and attributes its authorship to Rawls.<sup>3</sup> Rawls himself, while in his interpretivist phase, calls his theory of justice "procedurally **neutral**".<sup>4</sup> Bruce Ackerman, whose view on the proper limits of the moral authority and

responsibility of the state is similar to those discussed in Part II, claims that he has helped "popularize the notion that something called **Neutrality** [is] at the heart of contemporary liberalism."<sup>a</sup>

In my view both the foundationalist and the interpretivist advocates of what I have called "state constraint" adopted the notion of neutrality for the same reason. They saw in its intuitive appeal a way to make their position more intelligible, more accessible to more people and thus to cover the distance between theory (the foundations or the abstract principles derived through social interpretation) and policy quickly and effortlessly. One could, therefore, say that the adoption of "neutrality" discloses the political ambition of the aforementioned philosophers and shows that popularizing their positions and turning them into forceful ideologies is for them a priority that ranks second only to the discovery of philosophical truth.

Now, as I mentioned in chapter 1, "neutrality" has a broad semantic range as it appears in many different contexts. The fact that it also belongs to the language of international relations makes its meaning even more difficult to determine (see Chapter 1). Interpretivists, being more aware of these facts, are more reluctant to adopt "neutrality". The

---

<sup>a</sup>Bruce Ackerman, "Neutralities" in Douglass, Mara, Richardson (eds.), Liberalism and the Good. (London: Routledge, 1990), p. 29. From a meta-theoretical point of view "liberalism" means either a particular interpretation of western democratic tradition or a particular foundationalist position. In politics this distinction is not significant, but for the philosopher who seeks to defend or criticize a particular (either foundationalist or interpretivist) position it makes a world of difference. (See also 6.3).

interpretivist Rawls is a case in point. Although he does not abandon "neutrality" he states that "the term...is unfortunate; some of its connotations are highly misleading, while others suggest altogether impracticable principles...But with due precautions taken and using it only as a stage piece, as it were, we may clarify..." the state constraint position.<sup>5</sup>

Soon foundationalist advocates of state constraint became equally cautious. In his Liberalism, Community and Culture Kymlicka hardly ever uses the term. In his recent writings Ackerman appears apologetic for adopting neutrality in his Social Justice in the Liberal State without making important qualifications.<sup>6</sup> Dworkin shows caution in using the notion of "state independence" in conjunction with "neutrality".<sup>7</sup> Nevertheless, none of the neutralists or their critics ever abandoned the idea of neutrality altogether. This means that it is impossible to quote and discuss the arguments of the participants in the so-called "neutrality debate" without re-inserting "neutrality" in the place of "state constraint". Of course, this is not to say that a careful application of the notion of neutrality is preferable to its replacement. Neutrality is a term whose particularism is unsuitable for the foundationalist and whose meaning is too unsettled for the interpretivist to turn into a normative principle.

It is for these reasons that many philosophers have attempted to map out the different possible meanings of "state neutrality". However, the distinctions and clarifications offered do not include the vital meta-theoretical distinctions of Chapter 2. My main objective in this chapter is, therefore,

to amend the map of all the different notions of "state neutrality" that are present in the literature by incorporating into it the distinction between interpretivist and foundationalist state constraint as well as the distinction between philosophical and political argument. This will help to clarify the positions of both the advocates and the critics of neutrality.

## 6.1 Neutralities

The most elaborate account of the different possible meanings of state neutrality is proffered by Rawls in "The Priority of Right and Ideas of the Good".<sup>8</sup> Rawls distinguishes between two main state neutralities: First, what he calls "procedural neutrality" (what Kymlicka more appropriately dubs "justificatory neutrality"<sup>9</sup>), according to which the institutional structure and government policy are justified "neutrally" and, second, what he calls "neutrality of aim" according to which "the aims of the basic institutions and public policy are neutral."<sup>10</sup> With respect to the former, Rawls states that "neutral justification" can mean either justification "by appeal to neutral values", or justification "without appeal to any moral values at all", or justification by appeal to values underlying the principles of free rational discussion.<sup>11</sup> Given that the last one of these three possible justifications can easily be incorporated into the first one, we are left with two alternatives: justification in terms of something called "**neutral** moral principles", and what we could call "amoral justification".

A couple of questions arise immediately: What exactly is a neutral moral value or principle? And how is it possible to offer an amoral justification for public policy or for a particular arrangement of the institutional structure? Rawls' answer to the first question is that neutral are "...values such as impartiality, consistency in the application of general principles to all reasonably related cases (compare the judicial principle that cases similar in relevant respects are to be treated similarly), equal opportunity for the contending parties to present their cases and the like."<sup>12</sup>

For a more complete answer one has to make use of the distinction between foundationalist and interpretivist philosophies of philosophy. Rawls' answer is offered during his interpretivist period. This is evident in the passage quoted above: Certain values are neutral because they are tacitly acknowledged as such by being characteristic of particular social practices or institutions (e.g. the judiciary) in democratic societies. The constructivist interpretivist in particular holds that we should regard these values as elements of a specifically political, characteristically Western morality which is distinguishable from the non-political moral outlooks of much wider scope, namely philosophical truths, religious beliefs and conceptions of personal virtue or character.

The foundationalist neutralist's position is rather different. He also maintains that there are two different kinds of moralities and that only the values of a particular kind of morality are actually neutral. His taxonomy, however,

is different. On the one hand there is the morality which is derived from what is essentially human and on the other there are all those lower-order moral outlooks and visions of the good life that are contingent. The values and principles of the essentialist morality are neutral because they do not conflict with those of contingent moralities as they are not their alternatives. A person or group can hold any particular moral outlook or conception of the good without violating the principles of essentialist morality. It is the enforcement of the latter that justifies social policy and the institutional structure of society.

One could say that the essentialist morality is an overarching, higher-order system of values whereas the specifically Western political morality is a set of values adapted for a particular purpose, namely the interaction between people qua citizens. The end result, however, is the same in both cases. Both moralities regulate the pursuit of the various conceptions of the good and are regarded as compatible with them.

Now, Rawls' answer to the second question posed earlier is that an amoral justification, i.e. a justification which could also be regarded as neutral between different moral outlooks, "...seems impossible, since showing something justified [in our case social policy and the general institutional arrangement] appears to involve an appeal to some values..."<sup>13</sup> Most participants in the neutrality debate share this view. A notable exception is Raz, who criticizes amoral justification but treats it as a well-established

position. Since, however, Raz' account of different notions of state neutralities does not coincide with that of Rawls, I will turn to it in the next section.

To complete

Rawls' account of the different possible meanings of "state neutrality" we must now turn to "neutrality of aim", that is, to the view that the objectives of governmental policy and institutional operation must be neutral towards the differing conceptions of the good. For Rawls neutrality of aim has three different meanings. According to the first one, which we could call "neutrality of opportunity"<sup>b</sup>, the state "is to ensure for all citizens equal opportunity to advance any conception of the good they freely affirm".<sup>14</sup> According to the second meaning, neutrality of aim holds that the state is to do nothing intended to favour or promote any particular conception of the good over another, or to give special assistance to those who pursue it. A suitable title for this kind of neutrality is "neutrality of intentions".<sup>c</sup> The third meaning of neutrality of aim is what Rawls himself calls "neutrality of effect".<sup>d</sup> According to this, the state is to do nothing that makes it more likely that individuals will choose

---

<sup>b</sup>It is William Galston who coins this term in his Liberal Purposes, (Cambridge University Press: Cambridge, 1991), p. 100. Galston clearly attributes neutrality of opportunity to Rawls.

<sup>c</sup>The term is used by Susan Mendus in her Toleration and the Limits of Liberalism, (Macmillan: Hampshire, 1989), p. 130. Galston refers to neutrality of intentions as "neutrality of aim" without distinguishing between the different meanings of the latter.

<sup>d</sup>In the literature this also appears as "neutrality of influence", "neutrality of outcome" and "consequential neutrality".

one conception of the good rather than another unless steps are taken to cancel, or to compensate for, the effects of policies or institutional mechanisms that do this. When it comes to ethnic cultures this view often appears in a different, stricter form: the state is to take measures to maintain the level of support that each conception of the good enjoyed before all conceptions became the objects of individuals' choice within a competitive (market) environment.<sup>e</sup> Ethnic communities of immigrants often claim that neutrality involves giving them the means to avoid assimilation by the dominant culture and maintain roughly the same membership.<sup>f</sup>

It is important to clarify the relation between neutrality of aim and procedural neutrality and connect them with the positions discussed in the previous chapters. Does the one entail the other, or can each one hold independently? Richard Arneson thinks that the latter is true and offers the following two examples:

An example of a state policy that satisfies neutrality of procedure but not neutrality of aim would be a policy of state establishment of religion based not on the judgement that the favoured religion is intrinsically superior to its rivals but rather on the estimate that promoting one religion over its rivals will facilitate the maintenance of civil peace. An example of policy that satisfied neutrality of aim but not neutrality of procedure

---

<sup>e</sup>I take it that this stricter form is what Kymlicka has in mind in "Liberal Individualism and Liberal Neutrality" when he talks of "consequential neutrality". This point was made in the discussion of Kymlicka's foundationalism in Chapter 4.

<sup>f</sup>Or, more precisely, that their membership should depend only on the birth and death rate within the community.

would be a policy of broad religious toleration that aims to favour no religion over another and that is justified by appeal to the judgement that Quakerism is the true religion and among the tenets of Quakerism is the principle that there should be broad religious toleration and no tilting by the state in favour of any one religion.<sup>15</sup>

Two points should be made in connection with this view. First, it is evident that Arneson has in mind a particular neutrality of aim, namely neutrality of intention. Perhaps the narrowness of his view is determined by his choice of example, i.e. Quakerism. Unlike the distinctly political morality of Rawls' interpretivist phase, Quakerism is not associated with a scheme of distribution of resources and opportunities and, therefore, the claim that Quakerism upholds neutrality of opportunity or neutrality of effect would be absurd.

Second, as an example of neutral justification (i.e. procedural neutrality) Arneson offers justification in terms of a single, reputedly neutral, value, namely civil peace. Now it could, perhaps, be argued that civil peace is not neutral in the sense that it is not attributed the same importance by all the different conceptions of the good. Certain Christian sects (e.g. the Quakers) think of it as more important and are willing to make more compromises in its name than, say, Moslems do. Perhaps this is an effective argument against the view that single values can be used in neutral justifications of policies or institutional arrangements. This, however, is not the point I want to make here. Rather, the point is that in contrast with the Quakers and pacifists of Arneson's example, the philosophers discussed in the last two chapters

justify policies and institutional arrangements in terms of sets of principles or systems of values which they regard as higher-order (in the case of neo-Kantian foundationalism) or specialized (in the case of Rawlsian, interpretivist, distinctly political position) moralities rather than in terms of individual values.

So the conclusion we have to draw at this point is that according to the positions examined in the previous chapters the state is neutral neither because it is manned by people who subscribe to a tolerant conception of the good (and which remains but one conception among the many competing ones, e.g. Quakerism) nor because its functions are justified in terms of a value which is presumed to be common to all conceptions of the good (e.g. civil peace). Instead, it is neutral because it draws on a morality of a different kind, one which is in some sense compatible with all the differing conceptions of the good.

Having clarified the relation between procedural neutrality and neutrality of aim, it is important to examine whether the two state constraint positions (i.e. the foundationalist and the interpretivist) uphold the three different senses of neutrality of aim. First, are they both neutralities of opportunity? Rawls partly answers this question in his recent writings stating that his (interpretivist) position "...excludes the first meaning of neutrality of aim [i.e. neutrality of opportunity], for it allows only permissible conceptions (those that respect the principles of justice) to

be pursued." Nevertheless, he adds that "...that meaning can be amended to allow for this; as thus amended, the state is to secure equal opportunity to advance any permissible conception. In this case, depending on the meaning of equal opportunity, justice as fairness may be neutral in aim."<sup>16</sup>

It is clear that, in Rawls' view, subversive conceptions of the good are to be struck off the list of conceptions that the state allows its citizens to pursue without any kind of interference or restriction. But does this apply to the foundationalist reading of his position? As we saw in chapter 3 it is only "modes of conduct" that the state can prohibit or punish and not adherence to particular conceptions of the good. It appears, therefore, that Rawls' interpretivist state constraint position is less inclusive (of conceptions that are to be treated neutrally) than the foundationalist reading of his state constraint position. However, I believe that this is not the case. Although it is true that in some Western democratic societies certain moral/political conceptions<sup>g</sup> are regarded as unacceptable and that others are thought to be so radically opposed to the principles latent in political practice that cannot be regarded as valid choices of the free and equal members of a democratic society, such outlooks would not be regarded as conceptions of the good by either the neo-Kantian foundationalist or the Rawlsian constructivist interpretivist. This is because in both views conceptions of the good are not thought to be reducible to harmful activities

---

<sup>g</sup>Certain forms of organized racism, for instance the neo-Nazi political groups of Western Europe or the Ku Klux Klan in the U.S., are the obvious examples.

or subversive programmes and because both the neo-Kantian foundationalist and the Rawlsian interpretivist find that there is space for all conceptions of the good within a society governed by the principles of justice.<sup>16</sup> It seems, therefore, that Rawls' latest statement on the status of subversive conceptions is a direct interpretivist argument which is not fully in tune with his wider argument because it includes a different definition of "conception of the good".

However, one could claim that there is no real difference between discriminating against subversive conceptions of the good and prohibiting harmful modes of conduct. After all, what constitutes a harmful action depends either on the foundationalist's essentialist account (from which his harm principle is derived) or on the interpretivist's interpretation of the concept of harm latent in the particular political culture under investigation. Kymlicka, for instance, suggests that pornography constitutes harm because it contributes to the creation and propagation of stereotypes which set limits on the liberty and thus compromises the autonomous nature of women.<sup>17</sup> In similar fashion Raphael Cohen-Almagor avers that "psychological offence is on a par with harm."<sup>18</sup> Thus the neo-Nazis holding a non-violent march in a predominantly Jewish suburb of an American city are

---

<sup>16</sup>Of course there is a significant difference between the interpretivist and the foundationalist notion of a conception of the good: the Rawlsian interpretivist thinks of neo-Kantian foundationalism as yet another conception of the good.

considered to be violating the harm principle.<sup>1</sup> The implication of this view is that the bigots' (political) actions should not only be restricted when in violation of criminal law (e.g. when they can be classified as arson or assault and battery) but should be seen in the context of their declared subversive intentions. The question that arises here is whether we can accept this view and at the same time maintain that the bigots are as free as all others to pursue their conception of the good.

Oddly enough those authors who believe that bigots should be discriminated against or be altogether expelled from society acknowledge the significance of the distinction between subversive modes of conduct and conceptions of the good associated with such actions. For instance Cohen-Almagor states that "some people may adopt a conception that they see as a conception of the good, but that we regard as one of evil. If **its consequences are harm to others**, then we should not tolerate that conception."<sup>19</sup> This point is completed when Cohen-Almagor states that "fascists exclude themselves from liberal society **not because they hold undesirable beliefs**. Instead, the combination of holding illiberal beliefs **and** acting in accordance with them affords grounds for exclusion."<sup>20</sup> Taking these two statements in conjunction one concludes that it is only harmful actions that the state prohibits and punishes and not adherence to particular conceptions. Racists may receive welfare benefits, run for

---

<sup>1</sup>Cohen-Almagor refers to the Skokie incident which took place in 1977 in the town of Skokie of the Chicago greater area (The Boundaries of Liberty and Tolerance, pp. 132-147).

public office under a racist manifesto (take for example the British National Party and the Front Nationale in France), vote and publish their views provided they do not violate the harm principle of the neutral, higher-order morality. Like Rawls (see chapter 3) and Cohen-Almagor, Barry is officially in favour of discriminating against the bigots but on one occasion comes close to accepting the view that it is certain actions and not subscription to particular conceptions that warrant the restriction of the bigots' liberty. He states that "the only response worth making [towards the bigots] is to try to defeat them politically, and if necessary, seek to repress them by force."<sup>21</sup> In the first half of this sentence it is made clear that if the bigots are to be subdued this will have to be done in accordance with the principles and rules predicated of the basic institutional structure and applying to the adherents of all the differing conceptions of the good. In the second half, however, Barry states that "if necessary" the bigots must be dealt with as if they were not equal members of society. This is an idea that was discussed in the context of Rawls' foundationalism in Chapter 3 and again as part of his constructivist argument in Chapter 5. When the subversive elements become strong enough to jeopardize the smooth operation and existence of the just institutions their suppression is warranted. In terms of the neo-Kantian foundationalist argument this point could be interpreted in two ways. First, it could mean that the bigots' autonomous pursuit of their conception of the good ranks second to the proper function of the institutions which exist in order to

maintain the essential character of each individual member of society. The bigots are denied their liberty so that more people can pursue their conceptions of the good uncoerced, being the recipients of primary goods distributed so as to offset undeserved inequalities between individuals.<sup>j</sup>

As far as neo-Kantian foundationalism is concerned, there are two problems with this interpretation. First, it appears to be incoherent. How can foundationalist morality require the deprivation of certain individuals (e.g. those holding racist conceptions of the good) of those rights and goods which they are entitled to qua essential human beings? If the main foundationalist meta-theoretical assumption is that there are certain essentially human characteristics and that foundationalist morality (which is to govern the institutional structure of society) should treat all those who bear these characteristics as equals, arguing for the repression of bigots along foundationalist lines means regarding them as less than human, as less than autonomous (rational) human beings. Although this is exactly how some people regard members of neo-Nazi organizations in Western democratic societies, the implication for the foundationalist is that one of the essentially human characteristics is choosing from a limited class of conceptions of the good. In other words, subscribing to these "permissible" conceptions is part of

---

<sup>j</sup>"Undeserved" are inequalities that result from the arbitrary distribution of physical and intellectual attributes, social position and wealth that takes place at birth as well as those inequalities that cannot be traced back to an autonomous (rational) choice made after the rectification of puerperal inequalities.

human nature. I do not believe that neo-Kantian foundationalism (at least as it was presented in Chapter 3) includes this premise. Instead, what the neo-Kantian philosopher regards as essential to being human is the capacity to autonomously choose and pursue any conception of the good.

Barry and Rawls could respond that the security of the institutional arrangement takes priority over the maintenance of the bigots' essentially human character only in special cases, namely when the bigots become so influential or so powerful as to interfere with the operation of just institutions. It is only then that the liberty of the many (dependent on the existence of institutions governed by the principles of foundationalist morality) can be weighed against the liberty of the few. This is the second possible neo-Kantian interpretation of the quote by Barry. Those who subscribe to racist or chauvinist conceptions are indeed autonomous, rational human beings and as such they partake of the distribution of primary goods (whether these are just liberties, in accordance with the first foundationalist constraint argument, or liberties connected with opportunities, powers and income in accordance with the second foundationalist constraint argument). They are not to be persecuted or discriminated against for what they are (i.e. adherents of particular conceptions), but for acting in ways that cause harm to others. As Kymlicka argues in Liberalism, Community and Culture, no conception of the good is reducible to coercive, proselytizing modes of conduct. Surely it is

conceivable that a racist is against violence of any form or even against the very idea that the distribution of primary goods through the institutional structure should be colour-sensitive. What Barry and Rawls seem to be saying, however, is that if we think of the bigots as a homogeneous group aiming at the replacement (either by legal or by illegal means) of the institutions of justice with an alternative arrangement, justice requires that they are repressed before they achieve their aim. So bigots are respected qua essential characters until they come close to achieving their aim of toppling the institutions of foundationalist justice.

This point may not be fully compatible with the foundationalist state constraint position presented in Chapter 3 and may be taken to imply that foundationalist state constraint does not constitute neutrality of opportunity. However, the significance of this point is minimal because although the issue of subversion is a real one and has concerned a number of theorists and lawyers in Western democratic societies, what certain people seek to undermine in the actual world is not the neo-Kantian, foundationalist institutional arrangement since that has not been (fully) established. Juxtaposing an actual problem with a prescribed, foundationalist "to be" in order to determine the implications of the latter seems absurd. This distinction is, I believe, what Rawls has in mind when he states that in a just society people will have the propensity not to adopt subversive conceptions.<sup>22</sup> His point really is that subversion in a society whose basic structure is governed by neo-Kantian

principles of justice is an issue that we cannot (or need not) fully address.

Despite Rawls' recent claim that his interpretivist state constraint does not grant equal opportunities to the advocates of subversive doctrines, we have to conclude that his interpretivist constraint principle is as inclusive as the neo-Kantian foundationalist principle. As we saw in the previous chapter, Rawls finds that Western constitutional history shows that within the central range of the freedom of political speech there is no such thing as seditious libel, there are no prior constraints on the freedom of the press and, most importantly, there is full protection of revolutionary and subversive doctrines. It, therefore, appears that Rawls' recent rejection of this position is an attempt at appeasing his anti-neutralist critics.

Now, one might argue that the Rawlsian state constraint is not a neutrality of opportunity because the primary goods it distributes equally to all individuals are not equally useful in the pursuit of all conceptions of the good. For instance, those who wish to form a community and live apart from others in accordance with some religious truth or philosophical truth and those who choose the life of asceticism are arguably being discriminated against for having opted for these particular conceptions. The Rawlsian liberties, for example, enable one to forsake the demanding, restricted life of the religious community from the very early stages of initiation. But even those who remain within the community, who brave the adversities and come close to the

achievement of their chosen conception can easily opt out and shed all responsibility towards the rest of the community. This very fact alters the nature of the experience of living within such a community, distorts its meaning and reduces its significance. The individual himself, unlike those who choose non-communal conceptions, is hardly encouraged to persevere with the pursuit of his chosen goal. Relatedly, the ascetic is provided with the safety net of the Rawlsian distributive scheme which undermines his austere lifestyle and contradicts his severe self-denial.

This is an interesting argument against the view that foundationalist and interpretivist constraint are neutralities of opportunity and as it is closely connected with other lines of anti-neutralist criticism it will be more thoroughly discussed in the context of my critical examination of anti-neutralism in Section IV. For the time being it suffices to say that both the foundationalist and the interpretivist neutralist can easily respond to this point. The former can argue, as Kymlicka actually does (see 4.3), that the impossibility of exiting a community of adherents of a particular conception of the good (e.g. a religious community or a monastery) should never be regarded as an essential element or requirement of that conception. The revisability of all conceptions is derived from the essentialist account and is, therefore, built into the very definition of a conception of the good. In this view, religious fundamentalism is not an essential aspect of any religion; it is simply a policy pursued by some of its adherents, a policy which violates the

foundationalist principles predicated of the institutional structure of society. The fact that justice does not call for the (temporary) revocation of certain liberties in order for some to achieve their chosen conceptions (as they perceive them to be) does not mean that justice does not provide them with an equal opportunity to achieve these conceptions.

The interpretivist, on the other hand, may respond that in modern democratic societies individuals have a want for Rawlsian primary goods. Those individuals choosing conceptions like the monastic life or life in a commune (conceptions which are rather unpopular among the members of modern, Western societies) are not necessarily rejecting the primary goods as goods which ought to be distributed by the state. Their chosen conception is meaningful or valuable to them mainly because it involves a kind of withdrawal, a self-denial, the undertaking to carry burdens which they do not expect others to bear. Testing one's limits is linked with such choices and requires the availability of an opt-out clause which is in fact provided by the liberty principle.

The point can be extended to all communal conceptions of the good: in Western societies communal life has changed significantly. Walzer, who, although a critic of state neutrality, gives some credit to interpretivist state constraint, admits that "there cannot be much doubt that we (in the United States) live in a society where individuals are...continually separating from one another--continually in motion, often in solitary and apparently random motion..."<sup>23</sup> For the interpretivist advocate of state constraint this kind

of social mobility implies a widely shared want for Rawlsian primary goods. The just distribution of the latter does not impede the pursuit of communal conceptions of the good. In fact primary goods are as useful in the pursuit of communal conceptions (as these are now understood) as they are in the pursuit of individualist conceptions.

The question that I must now answer is whether the foundationalist and the interpretivist constraint positions are neutralities of outcome. In his recent writings Rawls states that

...it is surely impossible for the basic structure of a just constitutional regime not to have important effects and influences on which comprehensive doctrines endure and gain adherents over time, and it is futile to try to counteract these effects and influences, or even to ascertain for political purposes how pervasive they are. We must accept the facts of common-sense political sociology.<sup>24</sup>

This reference to "common-sense political sociology" is clear indication that Rawls has in mind his interpretivist position. The argument here is that in democratic, pluralist societies it is impossible to offset the effects of the market or of the institutions of democracy on certain conceptions of the good. Some of them will perish, others will remain on a subsistence level and some will become dominant. Any principle designed to offset the various effects and influences on them is foreign to the political culture of Western societies and would clearly conflict with the interpretivist liberty principle.

The same conclusion must be drawn about foundationalist state constraint. As was explained in Chapter 3, the neo-Kantian foundationalist starts with his essentialist account which includes the ideas of separateness, autonomy and equality. He then avers that in order for an individual to be treated justly he has to be treated qua essential character, that is as someone who exhibits the essentially human characteristics. This involves offsetting the effects of the initial--arbitrary--distribution of social position, wealth, talent and physical attributes (so that the essential equality is upheld) before holding individuals responsible for their choices (qua separate, autonomous beings). A market system operating after the rectification of the effects of the initial distribution is, according to this view, a just institution. As Kymlicka explains in "Liberal Individualism and Liberal Neutrality", conceptions of the good can be thought of as marketable goods. Individuals conceive of them or adopt them, develop them, advertise them and above all pursue them. But the pursuit and achievement of any conception costs money in accordance with the weight of the burden it imposes on others. Kymlicka makes clear that the pursuit of some conceptions will unavoidably be expensive and this is something that people are bound to take into account as they choose their conception (in conditions of "democratic equality"). This means that such conceptions will become unpopular and remain, so to speak, on the shelf. Given that initial, undeserved inequalities have been rectified, there is no reason whatsoever why the state should interfere to boost

those conceptions' chances of becoming popular or as popular as rival conceptions. Foundationalist state constraint is, in this sense, not a neutrality of outcome.

To illustrate the main points in this discussion let us take the example of the son of Bangladeshi immigrants who claims that neutrality requires that the British government takes measures to **protect** his chosen conception which is the culture of his parents' native country. The initial assumption here is that such a conception is quite expensive to pursue because it involves learning a foreign language, practising a religion which is not as well established as others and following unusual customs. If by "protection" the particular individual means the provision of particular primary goods (such as freedom of religion, freedom of movement, the opportunity to seek and to hold positions and to train for economic offices or vocations that guarantee income and wealth, the distribution of income so as to offset undeserved inequalities etc.) which will enable him to responsibly pursue his conception of the good, then what he really requests the British government to do is to uphold neutrality of opportunity. If, however, "protection of his conception" is taken to mean the adoption of policies that will give his conception the same chances of being chosen that the more established (and therefore cheaper) conceptions have, then his request is an altogether different one. He now requests that the state interferes with the operation of the market not in order to offer him as good a chance to achieve his chosen goals (whatever those may be) as others have, but in order to

promote a particular conception of the good until it is equally well established as the already popular conceptions. Clearly the adoption of such policies cannot be traced back to the neo-Kantian foundationalist's essentialist account. According to the latter, conceptions of the good are not the rather inflexible systems of beliefs and sets of priorities put forward by some authority or group (e.g. a religious organization or a political party). As Kymlicka points out, by choosing autonomously the individual elaborates and modifies a particular conception.<sup>25</sup> He does not accept it in toto or carries it in a "correct" way. The particular conception of our example does not have to be held in the same way the immigrant parents once held it nor does it contain some "essential" elements which have to be necessarily observed if he is to be regarded as a true adherent (say, holding Friday as a holiday dedicated to prayer). The "cultural marketplace", to use Kymlicka's term, will inevitably lead to changes and modifications. Making adjustments rather than holding on to reactionary interpretations is the only way a cultural minority can buttress up its conception in a competitive market environment. After all the willingness to adapt is assumed of all immigrants who willingly enter such an environment.

But what if the request for a neutrality of outcome is made by someone who has not yet chosen a conception of the good, by someone whose loyalties are not (yet) partisan? His point would be that letting economic considerations interfere with one's choice of conceptions of the good amounts to saying

that such conceptions are mere preferences with no moral value attached to them. Furthermore, economic considerations are determined by contingent factors (such as the fact that before Britain opened its doors to immigrants the major religion there was Christianity). Letting the economy determine which conceptions become or remain established is an acceptance of what is contingent rather than of what is essential and this is something that apparently contradicts foundationalist meta-theory.

This argument is based on a couple of misconceptions. First, the most basic morally relevant entity in neo-Kantian foundationalism is the individual and not any community of adherents to a conception of the good or the conception itself viewed as a complete, self-sufficient entity. Individuals do not get attached to conceptions of the good. The latter are either devised by the individual or taken as (often strong) suggestions. As individuals are essentially equal qua autonomous choice-makers who have an essential want for particular goods, the foundationalist higher-order morality requires that they are held responsible for the consequences of their choices when those are made after the rectification of undeserved inequalities (the accidents of birth). The market achieves exactly this: it holds individuals responsible for their choice of conceptions of the good. The differences in the popularity that the various differing conceptions attain, consequent upon the operation of the market, are irrelevant as far as the foundationalist higher-order morality

is concerned.<sup>k</sup>

Second, unlike the foundationalist morality, which is derived from the essences, conceptions of the good are contingent moralities and as such cannot be predicated of the institutional structure of society. This means that although they have moral significance for those who hold them, they are indeed on a par with mere preferences where justice is concerned.

The third and final question I have undertaken to discuss in the present section is whether the two different kinds of state constraint (i.e. foundationalist and interpretivist) are neutralities of intention; whether, that is, they hold that the state is not to take action intended to favour or promote a particular conception of the good over another or to give special assistance to those who pursue it. In his "The Priority of Right and Ideas of the Good", Rawls makes clear that his interpretivist position is not only a (qualified) neutrality of opportunity but that it also is a neutrality of intention "...in virtue of the features of a political conception: so long as the basic structure is regulated by such a view, its institutions are not intended to favour any

---

<sup>k</sup>In Chapter 4, I examined Kymlicka's claim that the individual must be provided with a rich cultural environment from which to choose his conception of the good. This I called the "cultural enrichment principle" and claimed that it is meant to be an amended version of the first foundationalist state constraint argument. So the state does have reason to interfere with the "cultural marketplace", but not in order to protect or support particular conceptions. The cultural enrichment principle involves the adoption of policies which do not refer to any single conception or group of conceptions. Tax exemptions for those who get involved with or contribute to any society or association pursuing a conception of the good is a case in point.

comprehensive doctrine."<sup>26</sup> The neo-Kantian foundationalist position is also a neutrality of intention in the sense that the state--governed by the morality deriving from what is essentially human--does not purposely promote particular conceptions of the good. It is, however, crucial to reiterate one crucial point: According to the interpretivist who, because of his specific meta-theoretical perspective, regards conceptions of the good as the tenets and dogmas of the actual groups, associations, religious organizations, political parties and schools of thought of his society, the neo-Kantian, foundationalist position is just another "comprehensive doctrine" and cannot, therefore, be neutral between other conceptions of the good. For his part the foundationalist would argue that it is the interpretivist's political morality that cannot be neutral since it is contingent and parochial.

The conclusion to be drawn from this discussion is that the two positions hold different views both of the neutral morality (the one to be predicated of the neutral state) and of what this morality treats neutrally (i.e. of the idea of a conception of the good). Although they can both be described as procedural neutralities and as neutralities of opportunity and intention, they cannot be defended or criticized along the same lines. Before I examine the anti-neutralist critique in the light of this clarification, I will focus on a different kind of neutrality which I will try to fit into the foundationalist/interpretivist distinction.

## 6.2 Welfarism: The Other Neutrality

As was mentioned during the discussion of procedural neutrality, Joseph Raz offers his own account of the different meanings of state neutrality (or what he calls "anti-perfectionism"), an account which differs significantly from that of Rawls. First of all he gives three "interpretations of political neutrality"<sup>1</sup> which correspond to the three Rawlsian neutralities of aim and are listed as follows:

1. No political action may be undertaken or justified on the ground that it promotes an ideal of the good nor on the ground that it enables individuals to pursue an ideal of the good.
2. No political action may be undertaken if it makes a difference to the likelihood that a person will endorse one conception of the good or another, or to his chances of realizing his conception of the good, unless other actions are undertaken which cancel out such effects.
3. One of the main goals of governmental authority, which is lexically prior to any other, is to ensure for all persons an equal ability to pursue in their lives and promote in their societies any ideal of the good of their choosing.<sup>27</sup>

It is quite obvious that the first of these meanings of state neutrality is identical with what Rawls calls "neutrality of intentions", the second with "neutrality of outcome" and the third with the Rawlsian "neutrality of opportunity". Of course this is hardly surprising. As Rawls himself makes clear in "The Priority of Right" his account of the different

---

<sup>1</sup>What Raz calls "political neutrality" is what I refer to as "state neutrality". In the next section I will argue that there is a conception of neutrality which is not philosophical at all. It is that conception that I will call "political neutrality".

neutralities "draws on Raz's formulations" so that he can respond to Raz' misrepresentation of his position.<sup>m</sup> What is, however, surprising is that Raz' equivalent to Rawls' "procedural neutrality" (to the idea that the institutional structure and government policy are justified neutrally, that is either in terms of neutral, moral values or in terms of no moral values at all) is a quite distinct idea which Raz calls "welfarism".<sup>n</sup> According to welfarism, "governments...[are] required to promote the goals that people have, without discrimination based on their moral merit."<sup>28</sup> This position can be regarded either as the crux of the single, neutral, higher-order morality which can govern the basic structure and provide the proper basis for governmental action (in the same way the neo-Kantian, foundationalist morality and the Rawlsian, distinctly political, interpretivist morality are regarded by their authors to be suitable for that purpose), or as an amoral principle adopted in the conviction that there are no valid moral values. Raz dubs the former interpretation of welfarism "political welfarism" and the latter "moral welfarism".<sup>29</sup>

Moral welfarism is, I believe, easy to dispense with. First of all, if it is indeed true that there are no valid

---

<sup>m</sup>John Rawls, "The Priority of Right and Ideas of the Good", p. 262. It should be noted that Raz understands Rawls to be an advocate of neutrality of outcome; Joseph Raz, The Morality of Freedom, pp. 117-18.

<sup>n</sup>The principle of welfarism as a basis for governmental action was first introduced by Brian Barry in Political Argument, (Harvester & Wheatsheaf: London, 1990), pp. 38-43. It is, however, the case that Barry uses the term "want-satisfaction" rather than "welfarism".

moral values or, to put it in terms of the foundationalist and interpretivist meta-theories, that there are no moral principles that can be derived from the essences or discovered lying under shared practices, there seems to be no reason why the satisfaction of people's preferences should be the state's main purpose. It is, in other words, impossible to link a radical, moral scepticism with the welfarist thesis without making certain essentialist (or interpretivist) assumptions. This means that moral welfarism boils down to political welfarism. Raz takes the opposite view when saying that "[al]though their political results are identical, the two doctrines represent radically different views of both morality and politics." To illustrate his point he offers the example of those "[m]oral welfarists who object to the proscription of so-called deviant sexual practices or of marijuana [and who] sometimes appeal to principles denying the state the right to enforce 'private morality', whereas what they mean is the denial of the state's right to enforce the wrong morality."<sup>30</sup> In this view the reason why the moral welfarist assumes the mantle of the political welfarist is because he does not need to show his true colours to promote his political agenda. According to Raz, appearing as a political welfarist "...will serve [the moral welfarist's] purpose just as well."<sup>31</sup> In my view moral welfarism is a political welfarist tactic, an argument formulated in order to gain the support of those who are reluctant to accept the foundationalist or interpretivist arguments in favour of welfarism. The reason why moral welfarists often appear as political welfarists is not because

they want to avoid unnecessary controversy, but because they are themselves political welfarists. Whether they choose to conceal their foundationalist or interpretivist positions in order to render the principle of welfarism more attractive to their audience is not relevant from a philosophical point of view.

What, therefore, remains to be examined in this section is political welfarism or--since there is no longer need to distinguish it from moral welfarism--simply "welfarism". As has already been mentioned, this can either be foundationalist or interpretivist. Foundationalist welfarism holds that the principle of welfarism should govern the institutional structure of society because it is derived from what is essentially human, whereas interpretivist welfarism affirms that the reason it should be predicated of the basic structure is that it is latent in the practices of Western democratic societies.

Although Raz does not distinguish between interpretivist and foundationalist welfarism, he does, briefly, discuss the main interpretivist welfarist argument according to which, "[t]he principles of representative government...entail a commitment by the democratically constituted authorities to act on welfarist considerations."<sup>32</sup> Raz does not make clear what he means by "principles of representative government", but we can safely assume that he means the election of representatives to a legislative assembly, where they are expected to vote on and pass legislation satisfying their constituents' preferences. Voting itself is thought to be

linked with the welfarist principle. Since it is impossible to satisfy all preferences, the government aims at satisfying as many preferences as possible or at achieving the greatest net balance of pleasure. Voting, either through representatives or directly, by means of referenda and Gallup polls, is instrumental in doing this. According to the argument, the operation of representative democracy in Western societies shows that welfarism is latent in their political culture.

In Justice as Impartiality Brian Barry explores this idea further. In his view, it is not voting per se that is relevant in determining whether welfarism (or what he prefers to call "want-satisfaction") is a tacitly acknowledged principle in Western democracies, but the **rationale** on which referenda and public-opinion polls are based in these societies. Do citizens vote, or respond to questions concerning, say, environmental issues posed to them by pollsters, as consumers (whose views about the importance of environmental protection projects are turned into costs with a notional cash value and compared with the costs of alternative projects which are not friendly to the environment) or do they feel the need to express their opinions on what is the right thing to do? Barry argues at some length that the latter is true. People who vote on issues in western societies assume that they are called upon to express views which are of moral significance and not to simply give information concerning their relevant levels of utility.<sup>33</sup> Barry's is, therefore, an anti-welfarist, interpretivist argument.

Like Barry, Raz also claims that welfarism is not latent

in Western, political practices. As he puts it, "[t]he principles of representative government guarantee some measure of control by the population over those in authority...[and] do not entail a commitment by the democratically constituted authorities to act on welfarist considerations alone."<sup>34</sup> Although this argument is rather elliptical, there is no doubt about its interpretivist character.

This interpretivist welfarism is, of course, distinguishable from **foundationalist** welfarism. The latter can take many different forms, but I will concentrate only on one which derives from two foundationalist assumptions. According to the first, what constitutes the basis of human nature is a certain sort of consciousness or experience, namely want-satisfaction. This is the satisfaction of having one's preferences and plans of life fulfilled. It follows that the foundationalist morality, which (axiomatically) is the only one that is fit to govern the basic structure of society, calls for the satisfaction of people's preferences without discrimination based on their (contingent) moral merit.

According to the second essentialist assumption of this foundationalist position, all individuals are equal in the sense that distinguishing characteristics (talents, physical attributes, sex etc.) are irrelevant where the foundationalist morality is concerned. One possible conclusion that can be drawn from this is that want-satisfaction is to be distributed in the most egalitarian way possible. This can be taken to mean that the aim of governmental policy should be the achievement of that average level of utility which shows the

highest level of those with the least.<sup>o</sup>

Whatever the actual formula to be prescribed, it is imperative that individuals' preferences and utility levels are made available to be processed. For that reason monitoring consumer choice, surveying public opinion and counting votes are essential for the operation of the institutional structure in accordance with the welfarist, foundationalist morality.

Now, it is not my intention to discuss the validity of the interpretivist and foundationalist welfarist positions. Rather, my point is that although welfarism appears in the literature as a neutralist position,<sup>p</sup> and although "welfarism" appears in discussions of Rawlsian justice,<sup>35</sup> the position examined in this section is not compatible either with neo-Kantian foundationalist neutrality or with Rawlsian interpretivist neutrality.

The differences between welfarist foundationalism and neo-Kantian foundationalism are rather obvious. First of all, the former, unlike the latter, does not call for the provision of means (e.g. the Rawlsian primary goods) that are regarded as useful in the pursuit of any chosen goal, but for direct, made-to-measure assistance. What the welfarist state

---

<sup>o</sup>A similar point is made by Barry in Justice as Impartiality, p.136. Nevertheless, Barry does not regard want-satisfaction as a foundationalist position and this is why he regards the introduction of egalitarian considerations into want-satisfaction theory as a flaw.

<sup>p</sup>Raz distinguishes between neutrality and welfarism but places both under the heading "anti-perfectionism" and talks of a close relation between them. Similarly, Barry regards want-satisfaction (the equivalent of Raz's "welfarism") as one of "two kinds of Neutrality", Justice as Impartiality, pp.139-45.

distributes is not a common currency which is to be used by all individuals in their pursuit of their conceptions of the good, but want-satisfaction, i.e. the satisfaction of achieving a held conception whatever that may be.

Second, in welfarist foundationalism the fact that all individuals are equal does not mean, as it does in neo-Kantian neutrality, that they have the right to choose and pursue their conceptions of the good from a position of near equality--where the accidents of birth have been offset by the proper distribution of primary goods--but that they all deserve to have their conceptions satisfied even if these conceptions were adopted in a state of inequality of resources, liberties and powers.

Third, according to welfarism, the individual qua essential character does not come by his conception of the good in a particular way. His conception is not necessarily the object of choice but could be constitutive of him, inculcated in him, or adopted simply for lack of alternatives.

The conclusion that can be drawn from these differences is that the two foundationalist positions are hardly compatible with each other. For the neo-Kantian foundationalist it is not only the process of choice-making that is essential to being human, but the pursuit of the chosen goal as well. In fact the two are inseparable since by pursuing his conception of the good one modifies or even transforms this conception. Both the choice and the pursuit of a conception are, in fact, connected with the idea of responsibility. Once all undeserved inequalities have been

rectified or compensated for, the individual is left to deal with the consequences of his choices. This is not the case with welfarist foundationalism where individuals are responsible only for providing accurate information concerning their preferences before this information is processed in accordance with a welfarist formula. It follows that any combination of the essentialist assumptions of these two foundationalist positions is bound to lead to problematic conclusions.<sup>4</sup>

Why is it, therefore, that, according to certain critics, Rawls should (given his essentialist account) prescribe the distribution of welfare rather than the distribution of primary goods? The argument goes along the following lines: since individuals are to be treated as essentially autonomous choice-makers, they can only be held responsible for their choices and not for accidents of birth or blows of bad luck. G.A. Cohen and Richard Arneson are among those who maintain that people who are inefficient in converting primary goods into chosen aims and whose inefficiency does not reflect any already made choices (e.g. being in physical pain caused by a genetically inherited illness), suffer an undeserved "welfare deficiency".<sup>5</sup> What is then required is the elimination of such

---

<sup>4</sup>The same can be said of interpretivist neutrality. Given that responsibility for one's choices is one of the main findings of the interpretivist Rawls, it follows that the two arguments of Chapter 5 cannot possibly incorporate welfarist elements.

<sup>5</sup>The term belongs to G.A. Cohen who amends Arneson's principle of "equal opportunity for welfare" in his "On the Currency of Egalitarian Justice", *Ethics*, 99 (1989), pp. 906-944. Cohen claims that Rawls undermines his egalitarianism as he fails to understand that his principles of justice

deficiencies. Although this view has been branded as "welfarism", it is quite obviously different from the welfarism described above. Whereas the welfarism Raz has in mind equates welfare with utility, the welfarism of Cohen simply appends to the neo-Kantian essentialist account a number of essential conditions. Barry, who notices the difference between the two welfarisms, suggests that these conditions include "the presence of adequate food, clothing and shelter, freedom from pain and suffering..."<sup>36</sup>

The conclusion to be drawn, therefore, is that utilitarian welfarism is a markedly different notion of state neutrality, one that is very rarely addressed by the participants in the neutrality debate. For this reason I will leave it here and move on to a non-philosophical notion of state neutrality, one that is not derived from an account of what is essentially human or discovered in the moral/political conceptions of modern democratic societies or thought to be an essential feature of Western political tradition, understood as an immutable, autonomous sphere of interrelated ideas (see Chapter 8). This state neutrality is distinctly political.

### **6.3 Liberalism and Political Neutrality**

Dworkin's neutralist position will serve as an introduction to the notion of political neutrality. Although Dworkin's argument is, as I will now argue, foundationalist in character, its formulation paves the way for the presentation

---

penalizes those who are blamelessly inefficient in their use of their share of primary goods (pp. 912-916).

of political neutrality.

Dworkin has written extensively on the issue of legal and social interpretation and for this reason his work is often regarded as a contribution to "hermeneutic political philosophy".<sup>37</sup> However, his general outlook, and more specifically his neutralist position, is hardly interpretivist in terms of the meta-theoretical distinction of Chapter 2. This is made clear in "Foundations of Liberal Equality" where he expresses his views on interpretivism

Can we adopt this idea, that principles have categorical force when they figure in the best interpretation of political history, to supply categorical force for a political conception of justice constructed so as to embody principles with that pedigree? I doubt it. No set of principles constructed by a philosopher can fit all parts of any community's traditions and history perfectly, because...the history of any community includes controversy as well as tradition. Two very different political conceptions, which would justify very different political decisions now, might each fit the record and rhetoric of a community's political history roughly equally well...We can only decide which principles are latent when we already have in hand some conception of justice whose categorical force we can defend in some other way, as not dependent on or derived from its congruence with the community's traditions.<sup>38</sup>

What Dworkin says here is that the philosopher must choose the interpretivist position which prescribes principles coinciding with, or being closer to, his foundationalist principles. The "conception of justice whose categorical force we can defend in some other way" is, therefore, none other than the Dworkinian foundationalist morality. An interpretation whose findings are evaluated in accordance with foundationalist

principles is not really an interpretivist philosophical position since interpretivist philosophy rejects foundationalist meta-theory. Hence, what Dworkin calls "the interpretation of political history" is actually a feature of his foundationalist argument which I will now examine.

As was argued in Chapter 3, the foundationalist's first move is to lay down the foundations, the essentially human characteristics which he has discovered. He then formulates moral/political principles which are presented as dictates of human nature (e.g. the Millian liberty principle) and, finally, offers more detailed political prescriptions (a political agenda) by elaborating his abstract moral principles. This is not to say that a foundationalist argument is a clear-cut, three-stage strategy, but that it is a process of specification, of determining the implications of having particular characteristics as human beings and that it is useful, from a methodological point of view, to think of it as having separate, well-defined stages.

As was mentioned in Chapter 3, the process of specification has to be completed without the introduction of elements which are not traceable to the foundations. This usually means that the more elaborate the foundationalist's prescriptions are, or rather, the longer the line connecting his prescriptions (his view of the just society) with the essentialist account, the more ad hoc they are bound to look and the more controversial the argument as a whole is bound to be. Dworkin's simple suggestion is that in order for the foundationalist to make sure that his detailed prescriptions

are indeed derived from the intermediate, abstract principles without the introduction of elements which cannot be traced back to the essentialist account, he should retrace his steps, that is start from the third and final stage of specification and move backwards to establish the connection with the intermediate stage.<sup>38</sup>

Although these meta-theoretical points are not explicitly made in "Foundations of Liberal Equality", they are actually applied in "Liberalism". Dworkin's objective there is to find the "constitutive element", the one abstract, general moral principle which constitutes "the nerve of liberalism".<sup>39</sup> Liberalism itself is taken to be a particular institutional arrangement, dubbed "the liberal settlement".

Not surprisingly Dworkin's proposed constitutive principle of liberalism is the principle of neutrality (i.e. the state constraint principle discussed in the previous chapters). Dworkin claims that the most fundamental assumption--shared by most of the foundationalist positions that are popular in Western democratic societies--is that individuals are essentially equals. This means that "the government [should] treat all those in its charge as equals."<sup>40</sup> Now, this point acquires different meaning within each one of the differing foundationalist views. According to the foundationalist position which yields the liberal arrangement, it takes the form of the principle of neutrality

---

<sup>38</sup>The connection between the abstract principles and the essentialist account does not interest Dworkin. He regards this early part of the foundationalist argument as rather straightforward.

which reads as follows:

Since the citizens of a society differ in their conceptions [of the good], the government does not treat them as equals if it prefers one conception to another, either because the officials believe that one is intrinsically superior, or because one is held by the more numerous or more powerful group.<sup>41</sup>

According to Dworkin liberalism endorses the market economy and representative democracy, regarding them both as basically egalitarian. Nevertheless, the principle of neutrality (as a specific interpretation of the requirements of equality) warrants a number of liberal principles which are meant to preclude or offset certain forms of favouritism resulting from the operation of these two institutions. The market economy tends to benefit those who are better endowed and those who belong to the upper socio-economic classes. Such people are more capable of achieving their conceptions of the good, or, to put it differently, their conceptions are bound to be better established than those of the untalented and those of the members of the lower classes. The principle of constraint is quite explicit here: the government should not prefer the conception "held by the...more powerful group" to the conceptions of other groups. Liberal endorsement of state intervention for the reduction of unemployment and the provision of services that would not otherwise be provided, as well as (liberal) support for racial equality and approval of governmental intervention to secure it, can be seen in this light.

Representative democracy can also favour some conceptions

of the good over others: A majority of legislators/representatives can pass morals laws branding certain conceptions of the good illegal not because their pursuit involves the violation of the rights of others, but because they are regarded as intrinsically inferior by the majority of citizens (or the majority of the representatives themselves). For this reason the state constraint principle warrants limitations on majority rule: The government should not favour a conception which is "held by the more numerous group" or a conception which is regarded as "intrinsically superior" by state officials. A number of liberal political positions is based on this point: rejection of all morals laws; opposition to the regulation of speech, sexual conduct and literature; imposition of procedural constraints and devices (e.g. rules against the admissibility of confessions) which reduce the extension of the provisions of criminal law to conduct of controversial morality and thus safeguard against the introduction of morals laws through the back door.

According to Dworkin, the principle of neutrality is indeed the intermediate principle which he set out to specify since it meets three basic criteria: it is "sufficiently well tied to the...liberal settlement," it is sufficiently detailed so as to discriminate liberalism from other egalitarian foundationalist moralities, and it is a comprehensive and frugal scheme that has great explanatory power and can, therefore, "provide...a fairer test of the thesis that it

precedes and survives particular settlements."<sup>1</sup>

Although the argument of Dworkin's "Liberalism" is a foundationalist argument, there is no denying that it is a rather unusual one. The idea that the foundationalist has somehow forgotten how he arrived at the various principles he prescribes and has to retrace his steps back to his first principles is rather paradoxical. The most likely explanation is that the argument is actually addressed to those "liberals", i.e. to those members of Western societies who accept most of these secondary principles (e.g. the regulation of the market and the rejection of morals laws), who do not know how these principles are related to one another. Dworkin argues that what links all of them together is the principle of neutrality.

Dworkin's view of neutrality as the constitutive element of liberalism should be contrasted with Brian Barry's view of neutrality as a mere support-seeking strategy. In his "How Not to Defend Liberal Institutions" Barry relates how liberals have been addressing different arguments to non-liberals in order "to show them that they too have good reason to support liberal institutions".<sup>42</sup> First, there was "the argument from social peace" inspired by and put forward after the European Wars of Religion. According to this, in order for social peace to be achieved, the adherents of any one conception of the

---

<sup>1</sup>Dworkin adds a fourth criterion that should be met by the intermediate, foundationalist principle: it must yield "...positions that it makes sense to suppose might be constitutive of political programs for people in our culture" (A Matter of Principle, pp. 186-7). This apparently confuses the argument and obscures the distinction between foundationalism and interpretivism.

good, and more particularly of any one religious dogma, have to adopt a policy of tolerance towards the heterodox members of society. According to Barry this argument fails because it assumes that all factions attribute a high value to social peace and that, therefore, society is already secularized.

Second, there was "the argument from prudence" according to which the adherents of a single religious dogma have to be tolerant towards the heterodox for fear that after pursuing an open conflict they are defeated and have an alien set of beliefs imposed on them. The argument clearly suggests the adoption of a risk-minimizing strategy and its weakness is, as Barry notes, that it assumes the existence of a balance of force within society. Liberalism's concern for minorities is, therefore, not consonant with this argument.

A third argument is the Lockean "argument from inefficacy" according to which religious persecution is pointless because religious belief cannot be coerced but can only be reached from within. In order to extend the force of the argument over all kinds of conceptions of the good and thus bring it up to date, we would have to say that conceptions of the good are only held when they are "made one's own", when they are adopted and pursued voluntarily and out of conviction. Barry points out that the adherents of many conceptions of the good do not share this view; instead they regard coercion as an effective means of bringing conviction (or religious belief) into line with performance. Such people cannot possibly be convinced in this way to adopt a more tolerant posture.

For Barry the principle of neutrality (state constraint) belongs to this long tradition of arguments "...that can be addressed to non-liberals to induce them to support liberal institutions while retaining intact their non-liberal attitudes."<sup>43</sup> In his view, such arguments are support-seeking (or, more accurately, consent-seeking) strategies. The political party upholding the liberal agenda promises to refrain from favouring one conception of the good over another so as to gain the support of the adherents of all conceptions of the good--including the support of the dogmatists. Barry avers that this policy is bound to fail because no dogmatist would ever endorse the liberal agenda. Islamists, for instance, would never subscribe to the "be and let be" message implicit in the principle and would accept nothing less than a Muslim organization of society which would, at the very least, involve making the teaching of Muslim customs and values part of public schools' curriculum, banning sexual literature, making illegal certain forms of sexual conduct and limiting the rights and opportunities of women. For Barry, such an arrangement is in direct conflict with the liberal social blueprint and so pursuit of neutrality seems to be the equivalent of unilateral disarmament on the part of liberalism.

Barry claims that there are clear indications that non-liberals will not hesitate to take advantage of this situation. As he puts it, "...the evidence is there for all to see. The churches that have made the biggest concessions to liberalism are losing ground while those that strain credulity

the most gain the most adherents....The popular press in Britain...reinforces and makes respectable the more atavistic impulses of its readers. The idea is gaining ground that schools should teach what parents want taught. And so on."<sup>44</sup> Liberals themselves are, therefore, bound to denounce neutrality on the basis that it is a potentially disastrous vote-winning strategy.

If neutrality is indeed a reward that non-liberals are promised for supporting the liberal agenda (or, in simpler terms, for voting liberal) and if the non-liberals' intentions are those described above, one has to accept Barry's conclusion. It should, however, be noted that in "How Not to Defend Liberal Institutions" Barry is concerned with the practicalities, the actual chances, of getting from "what is" to the liberal "to be". The kind of neutrality he criticizes is, therefore, different from those of chapters 3, 4 and 5. It is a **political**, rather than a philosophical, neutrality.

Political neutrality need not necessarily be a dictate or a corollary of the foundationalist, higher-order morality or a principle latent in moral and political practices. It can be a means of reaching the social arrangement dictated by a particular foundationalist or interpretivist position, a means of moving from the existing, not fully realized liberal arrangement to the one prescribed by political philosophy. This means that it may be external to the philosophical argument itself.<sup>45</sup> What is certain is that it is a concern of

---

<sup>44</sup>For example, it may be to liberalism what the dictatorship of the proletariat is to Marxist communism.

those statesmen who live in an imperfect world (where the institutions of justice are not yet fully in place) and who subscribe to neo-Kantian foundationalism or to Rawlsian interpretivism.

It is, of course, the case that both "philosophical neutrality" and "political neutrality" refer to the same principle of state constraint. It is, however, clear that the former applies to a society where the institutions of (either foundationalist or interpretive) justice are fully operative whereas the latter does not. Barry could be interpreted as saying that although neutrality may be the "nerve of liberalism", as Dworkin suggests, it is a bad policy in the sense that if applied in modern democratic societies, it will fail to bring about the neutral, liberal, social arrangement prescribed by liberal philosophy.

To avoid any misunderstandings, I should make clear that both neo-Kantian foundationalism and Rawlsian interpretivism assume that the higher-order, neutral morality is acceptable, or can be made acceptable, to a large majority of citizens. However, the need to make the philosophical prescriptions acceptable to a large majority of citizens does not affect the content of these prescriptions. The philosopher does not prescribe policies and arrangements which--he hopes--will be acceptable to the adherents of all conceptions of the good nor does he revise them when they are rejected as non-neutral. Any arguments, except his own, that are offered in justification of his prescriptions in order to make a particular political agenda more popular are political.

In his latest contribution to the neutrality debate Barry adopts the view that foundationalist neutrality is not really neutral. As he puts it, "...my object...is to seek an answer to... [the] charge that justice as impartiality is a fraud because it purports to rest on neutral foundations while in reality resting upon a distinctively liberal conception of the good...[T]he argument from autonomy does rest on a distinctively liberal conception of the good."<sup>45</sup> This means that Barry, once again, treats "neutrality" as a political term. Only **political neutrality** (as defined above) is relevant in the neutrality debate. Whether liberals like to think of liberalism as a neutral, higher-order conception that is compatible with all other conceptions is not relevant in what is a distinctively political debate. As in "How Not to Defend Liberal Institutions", Barry's concern is how the liberal agenda will be made acceptable to non-liberals. Advertising neo-Kantian foundationalism as neutral, that is using the philosophical argument itself, is a bad liberal strategy. But now Barry has come to think that the intuitive appeal of "neutrality" can do the trick. Political neutrality should not, therefore, be abandoned. Instead, it should be differentiated from foundationalist neutrality. The latter may be the way in which some liberals perceive of their (higher-order) conception of the good but it is useless as a support-winning position. It is not, however, considered as dangerous or self-defeating as it was argued in "How Not to Defend" and this is because it is now **not** taken to be a promise of unilateral disarmament in exchange for votes, but simply a

philosophical argument which, like its three predecessors, is lacking in political conviction. This means that although Barry has softened his position towards the presentation of neo-Kantian foundationalism as a neutral position, he still rejects it for being politically ineffective.

But the connection with his previous statement on neutrality is not the core of Barry's recent argument. The main point he now makes is that neutral can only be the constitutional framework which would be decided upon by individuals who are moved by the "agreement motive", that is by the desire to "live in a society whose members all freely accept its rules of justice and its major institutions...",<sup>46</sup> and are sceptical in the sense that they deny "the legitimacy of certainty".<sup>47</sup> Barry argues at length that for individuals to decide on a neutral arrangement it is essential that they bear both of these characteristics.

Is this a political argument? Is it offered in order to induce non-liberals to endorse the liberal agenda? The answer is No. Barry's clearly is a full-fledged interpretivist position. His point is not that the adoption of the idea of neutrality will help popularize the liberal institutional arrangement or that the liberal argument (if not the liberal institutional arrangement itself) should be modified to accommodate the adherents of all conceptions of the good, but that the two conditions stated above (i.e. the agreement motive and scepticism) are actually present in democratic societies. This is not to say that people there are self-declared sceptics who want to reach agreement on basic,

political principles. Rather, the point is that there is evidence in support of the view that the desire to justify one's actions in terms which are intelligible and acceptable to others and the rejection of the legitimacy of certainty are implicit in moral and political practice in Western democratic societies.<sup>v</sup>

The crux of Barry's recent argument is that the interaction between these two interpretivist findings leads to the idea of state neutrality. This means that the argument is a constructivist one. In the same way that Rawls puts together a number of interpretivist elements (i.e. the findings of separate interpretivist projects) in order to determine what principles are suitable for Western democratic societies (see 5.2), Barry combines his two interpretivist findings and arrives at the principle of neutrality. In his view, given

---

<sup>v</sup>I have, admittedly, removed the universalist overtones in Barry's argument. Barry seems to be suggesting that the desire to justify one's actions to others and the kind of scepticism that, as he says, underlies the rejection of the legitimacy of certainty, have been present in every historical society. For instance, he avers that religious scepticism has always been widespread and claims that "...the sheer weight of the evidence in favour of [religious] scepticism seems overwhelming" (Justice as Impartiality, p. 171). In my view, Barry's argument can only stand as an interpretivist one. The two elements he identifies are closely related to the ideas of public accountability and moral scepticism that are prevalent in Western democracies. Given that the examples of religious scepticism he offers (p. 170) are basically drawn from Western largely secular societies (e.g. the U.S. and 19th century Germany and Austria-Hungary), it appears that he would not reject the particularist interpretation of his argument.

There is, however, a further indication that Barry's argument is interpretivist in character: he accepts constitutional neutrality alone and maintains that at the legislative level the value of certain kinds of conceptions of the good should be debated on and put to the vote. This is a view that is widely shared in Western democracies and one which an interpretivist would hardly ever overlook.

that the desire to justify one's actions to others and the rejection of the legitimacy of certainty are both widely shared in Western democracies, one must conclude that no particular conception of the good should be built into their constitution.

Barry avers that neither one of these two elements can by itself lead to the principle of neutrality. This is a valid point: one may satisfy one's desire for justifying one's actions to others by justifying them in terms of one's particular conception of the good; justification does not mean "neutral justification" unless it is paired with scepticism. The latter does not by itself entail neutrality since, in Barry's own words, "if there is nothing to choose between conceptions of the good, what objection could there be to the government's picking one (perhaps at random, even) and enforcing it?"<sup>48</sup>

It is interesting to note that although this argument appears to be a complete reversal of Barry's earlier position on neutrality, his two arguments--the political argument of "How Not to Defend Liberal Institutions" and the philosophical argument of Justice as Impartiality--are actually compatible with each other. In the earlier argument Barry rejects foundationalist neutrality both as a philosophical position and as an argument designed to popularize the liberal agenda whereas in his recent statement he endorses interpretivist neutrality by means of a constructivist argument which is both simpler and more straightforward than Rawls'.

Now, the neutralists who are most aware of the difference

between philosophical neutrality and political neutrality are, I believe, Bruce Ackerman and Charles Larmore. Ackerman understands that in order to make liberal (neutral) institutions acceptable, the philosopher (who is either a neo-Kantian foundationalist or an interpretivist neutralist) should not simply present his philosophical argument. As he puts it,

[i]f I simply tried to tell you why I personally am a committed liberal, this would predictably divert some of you from arguments that you would find more compelling. Given this dilemma, it seems best to dispense with personal declarations of faith so as to better give you a sense of the different paths that can lead a thoughtful person to neutrality.<sup>49</sup>

Ackerman is not really saying that he is a sceptic (although he mistakenly suggests that the sceptic is bound to endorse neutrality). What he means is that a philosophical neutralist argument is unlikely to persuade non-liberals to accept state neutrality. His view is not that we cannot really know philosophical truth, but that knowledge of the truth does not make for persuasive arguments. What the philosopher needs to do in order to promote the establishment of his proposed institutional arrangement is to make appealing to the members of his audience. A single justification is hardly likely to do the job (especially if his audience is the kind of multi-racial, multi-cultural, multi-religious society that Ackerman is addressing). What he should, therefore, do is offer as many

disparate justifications of his position as he can muster.<sup>w</sup>

Ackerman purports to do exactly that. He claims that neutrality can be reached by many different routes. Neo-Kantian foundationalism (which he personally endorses) is just one of them. The "dialogic ideal", the values that underlie the principles of rational discussion, is another.<sup>50</sup> Scepticism, experimentation (the view that one has to work out one's conception of the good by oneself) and mistrust of those who hold political power also feature in Ackerman's list.<sup>51</sup> At times he even appears to be suggesting that an interpretivist investigation would also lead us to the principle of state neutrality.<sup>52</sup>

Ackerman's conclusion is that neutrality "...can be reached by countless pathways of argument coming from very different directions. As time passes, some paths are abandoned while others are worn smooth..."<sup>53</sup> Nevertheless, the paths to neutrality he actually mentions are not diverse enough to render his argument persuasive. The dialogic ideal, and the notions of scepticism and experimentation are all neo-Kantian spin-offs.<sup>x</sup> Of course, this point has not gone unnoticed by

---

<sup>w</sup>I believe that Larmore makes this very point in "Political Liberalism", Political Theory, 18 (1990), p. 345, when he avers that "...the Kantian and Millian conceptions of liberalism are not adequate solutions to the political problem of reasonable disagreement about the good life. They have themselves become part of the problem."

<sup>x</sup>The strong connection between the dialogic ideal and the neo-Kantian notion of nationalism is revealed by Ackerman himself in "Why Dialogue?", p. 8. As for his argument from scepticism, it rests on a strong egalitarian premise (as is noted by Barry in Justice as Impartiality, p. 172). Finally, the strong connection between autonomy and the notion of experimentation is quite clear.

Ackerman's critics. As Benjamin Barber points out, Ackerman's argument lacks the creativity, variety, flexibility and inventiveness of a political argument and, therefore, fails to persuade.<sup>54</sup>

The same can probably be said of Larmore's political argument. Like Ackerman, Larmore is actually exploring the arguments that the liberal (neutralist) philosopher could address to the members of his diverse audience in order to persuade them to accept state neutrality. He rejects autonomy, scepticism and experimentation for being too closely connected with the neo-Kantian position and suggests that sympathy, the desire for civil peace and the notion of "equal respect for everyone" are the only considerations that can appeal to the liberal philosopher's audience.<sup>55</sup>

Whatever its shortcomings, the neutralist's political argument is bound to be or to include philosophical positions. Although the philosopher wants to exert influence, he cannot possibly abandon his philosophy altogether and replace it with the empathetic and affective idiom, which Barber finds indispensable to political argument, without ceasing to be a philosopher.

Having made all the necessary distinctions between the different meanings of state neutrality, it is now possible to address the critique of state neutrality.

#### Notes

1. David Richards, Sex, Drugs, Death and the Law. (Totowa, New Jersey: Rowman, 1982), pp. 51-2.

2. Ronald Dworkin, "Liberalism", in A Matter of Principle (Cambridge University Press: Cambridge, 1985), p.191.
3. Will Kymlicka, "Liberal Individualism and Liberal Neutrality" in Ethics, 99 (1989), pp. 883-905.
4. John Rawls, "The priority of Right and Ideas of the Good", Philosophy and Public Affairs, 17 (1988), p. 261.
5. Ibid., p. 260.
6. Bruce Ackerman, "Neutralities", in R. Douglass et. al. (eds.), Liberalism and the Good (Routledge: New York, 1990), p. 29.
7. Ronald Dworkin, supra note 2, p. 191.
8. John Rawls, supra note 4, pp. 262-2.
9. Will Kymlicka, supra note 3, p. 886.
10. Ibid., p. 262.
11. Ibid., p. 261.
12. Ibid., p. 261.
13. Ibid., p. 261.
14. John Rawls, supra note 4, p. 262.
15. Richard Arneson, "Neutrality and Utility", Canadian Journal of Philosophy, 20 (1990), pp. 218-9.
16. John Rawls, supra note 4, p. 262.
17. Will Kymlicka, Liberalism, Community and Culture (Clarendon: Oxford, 1991), p. 167.
18. Raphael Cohen Almagor, The Boundaries of Liberty and Tolerance, (Florida University Press: Gainesville, 1994), p. 128.
19. Ibid., p. 49. (Emphasis mine).
20. Ibid. p. 85. (Emphasis mine).
21. Brian Barry, Justice as Impartiality, (Oxford University Press: Oxford, 1985), p. 169.
22. John Rawls, A Theory of Justice
23. Michael Walzer, "The Communitarian Critique of Liberalism", Political Theory, 18 (1990), p. 11.
24. John Rawls, supra note 4, p. 262.

25. Will Kymlicka, supra note 17, p. 168.
26. John Rawls, supra note 4, p. 262.
27. Joseph Raz, The Morality of Freedom, (Oxford: Clarendon Press, 1986), pp. 114-15.
28. Ibid., p.136.
29. Ibid., pp. 136-39.
30. Ibid., p. 139.
31. Ibid., p. 139.
32. Ibid., p. 139.
33. Brian Barry, supra note 21, pp. 154-59.
34. Joseph Raz, supra note 27, p. 139.
35. See, for example, G.A. Cohen, "The Currency of Egalitarian Justice", Ethics, 99 (1989), pp. 906-944; Ronald Dworkin, "What is Equality? Part I: Equality of Welfare", Philosophy and Public Affairs, 10 (1981), pp. 185-246.
36. Brian Barry, supra note 21, p. 140.
37. See for example Georgia Warnke, Justice and Interpretation (Polity Press: Cambridge, 1992), pp. 1-13, 62-87 and Stephen Guest, Ronald Dworkin (Edinburgh University Press: Edinburgh, 1992), pp. 20-42.
38. Ronald Dworkin, "Foundations of Liberal Equality", Tanner Lectures on human Rights, XI, 1990, pp. 32-34.
39. Ronald Dworkin, "Liberalism", A Matter of Principle (Cambridge: Cambridge University Press, 1985), p. 183.
40. Ibid., p.190.
41. Ibid., p.191.
42. Brian Barry, "How Not to Defend Liberal Institutions", in R. Douglass et. al. (eds.), Liberalism and the Good. (Routledge: London, 1990), p. 46?
43. Ibid., p. 49.
44. Ibid., p. 57.
45. Brian Barry, supra note 21 , p. 183.
46. Ibid., p. 164.

47. Ibid., p. 171.
48. Ibid., p. 173.
49. Bruce Ackerman, Liberal Justice in the Liberal State (New Haven: Yale University Press, 1980), p. 360.
50. Bruce Ackerman, "Why Dialogue?" The Journal of Philosophy, 86 (1989), p. 7.
51. Ibid., pp. 11-12.
52. Ibid., p. 12; "Why Dialogue?" The Journal of Philosophy, 56 (1989), p. 21.
53. Bruce Ackerman, supra note 49, p. 12.
54. Benjamin Barber, The Conquest of Politics (Princeton University Press: Princeton, 1989), pp. 120-151.
55. Charles Larmore, Patterns of Moral Complexity (Cambridge University Press: Cambridge, 1987), pp. 50-69.

## CHAPTER 7

### The Critique of Neutrality

My main objective in the present chapter is to make a more systematic presentation of those arguments which reject the idea of state neutrality and argue against the adoption of the principle of neutrality (state constraint), and to show that not all of them are aimed at the same target or meet with success.

Anti-neutralist arguments fall into four different lines of criticism. According to the first one, state neutrality is impracticable in the sense that it cannot be translated into workable principles and guidelines which can govern the institutional structure of society or set limits to governmental action. According to the second one, state neutrality is not actually "neutral" in that it either discriminates between different conceptions of the good or is derived from non-neutral values. The third general criticism holds that although state neutrality may be really neutral and even feasible it is bound to be very unpopular. Finally, the fourth criticism argues that the philosophical argument(s) justifying state neutrality are somehow flawed. Although there is considerable overlapping between these criticisms I will now approach them separately and examine how each fares against the neutralist positions examined above.

## 7.1 The Impracticability of Neutrality

The view of neutrality as impossible to implement is quite widely held. As we saw in Chapter 3, John Danley claims that the main implication of neutrality is the assignment of different rights to the adherents of different conceptions of the good which amounts to the creation of "a crazy quilt of different jurisdictions and...a jumble of different sets of rights."<sup>1</sup> Such an arrangement is not simply difficult to supervise and finance; it is also virtually unfeasible. As we have seen, Danley equates "conceptions of the good" with "cultural communities", but as has been mentioned, the proper meaning of "a conception of the good" is much wider and includes preferences, lifestyles and religious beliefs which can be seen as independent of any particular ethnic or cultural identities. It follows, therefore, that if Danley's "neutrality between cultures" is too expensive or too complicated to implement, a more complex "neutrality between all conceptions of the good" is totally inconceivable.<sup>a</sup>

It is, I believe, quite obvious that this line of criticism assumes that state neutrality is a neutrality of outcome whose aim is to make sure that market forces do not affect the popularity of the differing conceptions of the

---

<sup>a</sup>It is interesting to note that William Galston takes the exact opposite view when he states that "[t]he costs of cooperation under common rules with individuals who differ radically from me may appear prohibitive...It may be rational for me to prefer a multiplicity of separate homogeneous communities." (Liberal Purposes, pp. 147-148). The point here is not that state neutrality is a neutrality of outcome, a position which is totally impracticable, but that it is an expensive alternative to neutrality of outcome. The conclusion, however, is the same in both cases.

good. The latter are themselves viewed as fixed sets of principles and moral or aesthetic priorities which cannot be revised without being abandoned. The groups of adherents are also assumed to be numerically fixed. Any changes in the number of adherents are, therefore, thought to be in breach of neutrality. These assumptions lead to an idea of neutrality which is very different from the neo-Kantian foundationalist and the Rawlsian interpretivist positions examined above.

Ackerman responds to this particular line of anti-neutralist criticism in a similar manner. As he puts it, if neutrality is understood "...as a criterion...which seeks to judge political decisions by evaluating the **outcomes** they produce,...[n]eutrality seems a conceptual nonstarter... [M]y response is to deny that neutrality is a way of directly assessing consequences."<sup>2</sup>

The conclusion to be drawn from these points is that those who think of neutrality as inapplicable have in mind some kind of neutrality of outcome and, therefore, their criticisms can safely be dismissed as irrelevant. Nevertheless, Susan Mendus argues that neutrality of intention, too, is inapplicable in the sense that it can never be monitored or clearly established. As she puts it, "it is notoriously difficult to establish what the motivation is behind any particular piece of legislation and, given this fact, neutrality in motivation may simply be the fig leaf with which to disguise antagonism towards a particular group or groups. Ways of life may wither away and die not because they are clearly and overtly legislated against, but as a result of

the unintended, though foreseeable, consequences of some kinds of legislation."<sup>3</sup> Since both the neo-Kantian, foundationalist neutrality and the Rawlsian interpretivist neutrality are neutralities of intention, Mendus' view appears to be a relevant, powerful criticism.

Rawls and Barry sidestep this criticism by claiming that neutrality does not always operate at the legislative level.<sup>4</sup> In their view, there are bound to be certain issues (issues where individual rights are not in question) that cannot possibly be resolved without giving priority to some conception of the good over others. Although the procedure by which decisions on these issues are taken should conform with the principles of justice, a particular conception will, in the end, be promoted. Such issues are "...statutes protecting the environment and controlling pollution; establishing national parks and preserving wilderness areas and animal and plant species; and laying aside funds for museums and the arts."<sup>5</sup> To say that neo-Kantian foundationalist justice, or Rawlsian interpretivist justice, provides an answer to the question of whether a specific museum or theatre company should be publicly funded or that it determines the appropriate level of funding is absurd. A decision can only be reached in terms of a particular (lower-order) conception of the good (e.g. a conception which values highly Opera or Shakespearean theatre). Of course, any funds that go into such projects have to be raised in accordance with the principles

of justice.<sup>b</sup>

The upshot of this argument is that neutrality of intention is not elusive or impossible to confirm. The neutralist cannot suggest that in a just society all kinds of policies are adopted with neutral intent. The very nature of certain issues makes the promotion of some conceptions unavoidable. However, at the constitutional level neutrality of intention cannot be doubted. Constitutional principles (whether they derive from neo-Kantian foundationalism, or draw on principles implicit in a specifically political conception that is prevalent in modern democratic societies) are not designed to promote particular conceptions of the good.

## 7.2 Neutrality as a Fraud

Let us now come to the second main line of anti-neutralist criticism according to which state neutrality (as defined by the neutralists) may indeed be feasible, but is not actually "neutral" between the differing conceptions of the good.

This criticism is almost exclusively directed against foundationalist neutrality. Consider for example Barry's view of neo-Kantian neutrality. As he puts it, for the neo-Kantian foundationalist "[o]nly conceptions that have the right origins--those that have come about in ways that meet the criteria for self-determined belief--can form the basis for

---

<sup>b</sup>To this Barry adds that "the procedure by which the decision is taken should be fair" (Justice as Impartiality, p. 145). The point is not adequately explained and given that he states that state neutrality is incomplete as a decision rule (Ibid., p. 145), it feels rather redundant.

activity that has value. It is therefore unlikely that the good as autonomy will be advanced by distributing resources in a way that takes no account of the autonomous or non-autonomous origins of people's substantive conceptions of the good."<sup>6</sup> For this reason, Barry concludes, neo-Kantian foundationalism is not neutral.

Now, it seems to me that Barry's point does not really show that the neo-Kantian higher-order conception is non-neutral between lower-order conceptions of the good. Although the way in which a conception of the good is adopted is of great concern to the neo-Kantian foundationalist, its content is not. Any conception of the good may be pursued as long as it is chosen autonomously by the individual. It is in this sense that neo-Kantian foundationalism is neutral.

Of course, Barry is not alone in finding neo-Kantian foundationalism non-neutral. Mendus claims that although the neo-Kantian foundationalist argument is quite coherent, it is non-neutral and this is why it ultimately fails.<sup>7</sup> Galston makes the same point when he says that the more seriously neutralists take their commitment to autonomy/rationality, "...the more blurred becomes the line separating...[their proposed arrangement] from the tutelary, 'perfectionist' state committed to a fuller theory of the good."<sup>8</sup> The point that these critics are actually making is that the foundationalist's meaning of "neutrality" conflicts with the correct meaning of the word.

But what is this "correct" meaning? There are three possible answers to this question. According to the first

answer, which was discussed in Chapter 1, neutrality can either mean the lack of certain kinds of characteristics which would help us place someone or something in one among many specified classes or varieties, or it can mean taking neither side in a dispute, debate or difference of opinion. As I have mentioned (Chapter 1), these two definitions of "neutrality" are context-independent, but most theorists find them too sketchy to be of any significance. Be that as it may, they set standards which are definitely met by the neo-Kantian higher-order conception. Since the latter is only concerned with the conditions in which conceptions of the good are adopted, it lacks those characteristics that would make it one among the many, differing conceptions of the good.<sup>c</sup> Furthermore, it takes no sides in the conflict between rival conceptions of the good.<sup>d</sup> It is in this sense that it is a deemed a higher-order conception.

I believe that many of those critics who find neo-Kantian foundationalism non-neutral reject this definition of "neutrality". Instead they imply either that "neutrality" has a social meaning which can only be known through social interpretation and which sets a standard that neo-Kantian

---

<sup>c</sup>To put it differently it is not the object of one's choice. It is not adopted in order to satisfy a first-order or second-order desire or preference.

<sup>d</sup>This position can be maintained if we accept, as the neo-Kantian foundationalist does, that the way in which a conception of the good is adopted is not essential to that conception. If, in other words, coercive proselytism is not an essential feature of any conception of the good--which I believe is true--then neo-Kantian foundationalism cannot be thought to be taking sides in the conflict between rival conceptions.

foundationalism does not meet, or that neutrality depends solely on acceptability (i.e. that in order for an arrangement or policy to be neutral, it must be accepted as such by all competing groups of partisans) and that neo-Kantian foundationalism fails this test.

Surprisingly enough one possible response to such criticisms is provided by Barry himself who claims that "it is no argument against [neutrality] to say that [it] fail[s] to accord with some different conception of neutrality." If this statement is paired with the one which precedes it, according to which justice is neutral in the sense that "it proposes that all conceptions of the good should (**in a way defined by the theory itself**) be treated equally," we are bound to conclude that there is no need to compare the foundationalist's idea of neutrality with some other notion of neutrality which is external to "the (foundationalist) theory itself" in order to test its genuiness. The foundationalist morality is neutral because it is a higher-order morality which is compatible with all the differing conceptions of the good. "Compatibility" here is also defined by "the theory itself", as is the scope of all the differing conceptions of the good (in the sense that they are all regarded as lower-order conceptions and objects of individuals' autonomous choices). The real reasons why Barry rejects neo-Kantian, foundationalist neutrality are, as has been mentioned (see 6.3), that he is disinclined to accept foundationalist meta-theory and that he is interested in making his idea of neutrality politically attractive. To achieve this he has to

use as support-winning argument for the liberal cause some interpretivist position rather than an already widely criticised foundationalist view.

A second possible response is that the use of "neutrality" can be avoided without changing the actual principle itself which can be presented under the heading of "principle of state restraint". This was the argument in Part II. The idea of neutrality may be regarded by some as useful in explaining the neo-Kantian foundationalist position or the Rawlsian interpretivist position, but there is no reason why it should not be abandoned or replaced if the questions it raises are more than those it helps to answer.

Before closing this discussion on the second anti-neutralist criticism, I should briefly touch upon the issue of scepticism. In Part II I made clear that state neutrality does not derive from scepticism. However, certain critics have argued that the neutralist argument is an argument from scepticism and that, given that scepticism is not a higher-order conception of the good, state neutrality is not really neutral.<sup>c</sup>

---

<sup>c</sup>This is one of the anti-neutralist arguments proffered by Susan Mendus (Toleration and the Limits of Liberalism, pp. 75-79). In her view, historical evidence points to a strong connection between scepticism and a particular lifestyle, the life of ataraxia (imperturbability). This is a life of tranquillity that contrasts with the life of profound commitment or intense emotion. English seventeenth-century sceptics who embraced this way of life were "neutral" towards non-sceptics only insofar as their way of life was not threatened by the high emotion and disorder caused by their religious disputes. But being "mindful of the fact that certain practices were likely to cause civil strife, sceptics frequently took the course of intolerance and persecution" (p. 77). For Mendus this is enough proof that scepticism does not qualify as a neutral, higher-order conception and that the

It is important to note that this criticism is based on a serious misunderstanding. Scepticism not only fails to yield state neutrality, but fails to yield any principles that can govern the institutional structure of society. The sceptic is someone who believes that sure knowledge of philosophical truth, of the foundations, cannot be found. Assuming that he accepts that the principles that govern public life should be derived from the foundations, the sceptic cannot possibly prescribe any principles. For all he knows, the principles already governing the basic structure may be the ones that derive from the truth. Scepticism is, therefore, a reason for not prescribing principles and not the foundation of state neutrality. To put it in terms of the Platonic fable of the cave, the sceptic is one of the chained cave-dwellers who although aware that there is life outside the cave, in the true world of facts, he knows that no-one can escape from his bonds to acquire knowledge of it.

Neutralists often talk of scepticism, but this is not the foundationalist scepticism discussed here. Ackerman mentions scepticism as a possible route to neutrality, but he can hardly disguise the neo-Kantian foundationalism that underlies it. (See 6.3). Barry talks about a kind of scepticism which is an interpretivist finding and which does not lead to any principles unless it is combined with an other interpretivist finding, i.e. the desire to justify one's actions to others.

---

neutrality it endorses is really a fraud.

In fact the kind of scepticism discussed by Mendus is what Barry calls "latitudinarianism", a notion of scepticism which denies that all knowledge is impossible. (See below).

(See 6.3). Furthermore, he asserts that the notion of scepticism he identifies is a position lying somewhere between scepticism and dogmatism; although it rejects zealous dogmatism, it does not deny the possibility of any knowledge. In fact, he suggests that "latitudinarianism" is a better term for it, one that can help to distinguish it from extreme scepticism and is historically more accurate.<sup>9</sup>

Now, I have chosen not to address in this section the anti-neutralist argument that stems from the third definition of "neutrality" (i.e. from the one according to which neutral is what is accepted as such by the actual adherents of the differing conceptions of the good), because it falls into the third line of anti-neutralist criticism, a criticism that can be directed both against foundationalist and against interpretivist state neutrality.

### 7.3 The Unacceptability of Neutrality

One crucial question that seems to be of, at least some, concern to the participants in the neutrality debate is whether the idea of the neutral state is generally acceptable in Western democratic societies. Critics of state neutrality suggest that it is not and point to the fact that the adherents of many conceptions of the good (most notably, in Britain, members of the Muslim communities) reject it out of hand as discriminatory and detrimental to their ways of life.<sup>10</sup>

---

<sup>9</sup>William Galston explains the religious fundamentalists' position on state neutrality in Liberal Purposes (Cambridge University Press: Cambridge, 1990), pp. 130-131. As he puts it, from their perspective, state neutrality "...represents an

The significance of this point is not really clear.

Some critics seem to be saying that since not all sides accept the "neutral state" as actually neutral between the differing conceptions of the good, state neutrality is a fraud. It seems to me that if this definition of "neutrality" is accepted, then the term is clearly not suitable for a philosophical argument since the philosopher (whether he is a foundationalist or an interpretivist) does not formulate principles in order to accommodate the actual adherents of as many conceptions of the good as possible. The views that particular religious leaders, representatives of ethnic communities and spokesmen for cultural minorities hold on the genuineness of state neutrality is not taken into account by the neutralist philosopher in his formulation of the principles that should govern the basic structure of society. Therefore, shedding the term "neutrality", as I did throughout Part II, makes this criticism appear rather superficial.

However, the critic could suggest that the fact that people reject neutrality has serious implications both for foundationalist and for interpretivist neutrality. Starting with the former, he could argue that the idea of establishing (neutral) institutions which are not acceptable to a significant number of people seems to contradict the foundationalist's claim that the neutral state respects individuals' essentially autonomous nature. How can they be treated as autonomous if they live within an institutional

---

injunction to set aside God's word, the only source of salvation, in determining the principles of our public order."

structure which they categorically reject?

One possible response to this argument is to say that those individuals who find neutrality unacceptable do so because their autonomous nature is somehow impeded. Once the neutral institutions are in place, that is, once the proper conditions under which they can choose and develop their chosen conceptions are present, their attitude towards the neutral institutions will change. The fact, for instance, that in modern Britain a substantial minority rejects neutrality is irrelevant because the institutional structure there is not governed by the neo-Kantian, foundationalist morality. In this view, it is important to keep in mind that the present institutional arrangements in several modern democratic societies are not those derived from the neo-Kantian foundations and, therefore, their citizens' view of the neutral state is likely to be distorted.

Given, however, that actual social arrangements may be derivable from these foundations, or that there is a strong resemblance between the prescriptions of the neo-Kantian foundationalist and certain already established social arrangements, and as I do not wish to examine here the validity of such an argument, I believe that this response should be set aside.

According to an alternative response which, in my view, bears more credibility, a great many people oppose neutral institutions not because their autonomy is in some sense hindered, but because they are actually against the view that the adherents of conceptions which differ from their own

should be treated as autonomous choice-makers. The establishment of neutral institutions does not in any way compromise the autonomy of the anti-neutralists, since their autonomy is not dependent on their ability to coerce others into accepting their conception of the good.<sup>8</sup>

Now, with respect to interpretivist neutrality, the critic might argue that the rejection of neutral institutions by a substantial minority in Western democratic societies discredits the neutralist, interpretivist argument since the interpretivist philosopher himself acknowledges the importance that people in these societies give to public consensus. This means that the interpretivist findings themselves contradict the philosopher's distinction between people's opinions on matters of public morality, and the (philosophically relevant) principles which are implicit in moral and political practice.

As I argued in 5.2, this argument is based on a false assumption. The fact that the interpretivist neutralist prescribes a principle (i.e. that of state neutrality) which many actually reject, does not mean that he underestimates the value attributed to public consensus in democratic societies. Rather, he specifies the kind of consensus that is valued in Western societies and determines its institutional implications. In doing so, he pays closer attention to the undercurrents, to the principles and ideas implicit in

---

<sup>8</sup>The fact that Kantian autonomy implies acceptance of the foundationalist institutions is clearly not a problem. As was made clear in Chapter 3, the neo-Kantian notion of autonomy is narrower in scope.

practice, rather than to opinion polls or to the protests of the variously motivated leaders of religious or cultural minorities.

So the conclusion to be drawn from this discussion of the third line of anti-neutralist criticism is that the desirability or acceptability of state neutrality neither determines the content of the neutralist philosopher's prescriptions nor confirms the validity of his arguments. Both the foundationalist and the interpretivist neutralists have their reasons for believing that their prescriptions will be accepted by a significant majority of citizens. If some of them appear to be trying to accommodate every single dissenter, this is because, like Rawls, they are also concerned about the reception that their positions will get and the level of political support that they will be able to muster. To reiterate the point, the political philosopher's eagerness to gain popularity and exert influence, can sometimes obscure his account of the truth and its implications.

#### **7.4 Flaws in the Philosophical Neutralist Arguments**

It could be argued, that the three lines of criticisms examined so far are political in that they either point to difficulties in the achievement of a prescribed neutralist "to be", or find the use of the term "neutrality" by the neutralist philosopher as an attempt at misleading his audience, at making them accept his view as something that it is not. By contrast, the fourth line of criticism is purely

philosophical. It comprises arguments which, being sensitive to the meta-theoretical distinction between foundationalist and interpretivist neutrality, are concerned with the particulars of the philosophical neutralist arguments. Not surprisingly this includes arguments which are directed either against the foundationalist, neo-Kantian position or against Rawlsian interpretivism.

In the former category belong arguments which hold that the move from the essentialist account to the prescriptions is flawed or, to put it differently, that one cannot possibly get from the particular neo-Kantian essences to the principle of neutrality. If the foundationalist says that this move is possible it is either because somewhere along the line of specification he introduced further assumptions which are not related to his essentialist account, or because he overlooked some gap along the line of specification which is not possible to be bridged with the means available to him.

One such argument is proffered by William Galston who claims that the path from autonomy "...to negative freedom is obstructed, and a substantive doctrine of the good is...unsuited to serve as a foundation for negative freedom."<sup>10</sup> According to Galston the reason for this is twofold. First, autonomy (which for Galston is synonymous with "positive freedom") and negative freedom have different logics. One's autonomy can never conflict with that of another whereas negative freedom can. Second, the two of them lie in different spheres. Autonomy is metaphysical whereas negative freedom is not.<sup>11</sup> This means that the move from autonomy

(which is the basic element of the neo-Kantian's essentialist account) to the liberty principle constitutes a serious flaw in the foundationalist argument.

In my view, Galston is mistaken in assuming that the neo-Kantian starting-point is the Kantian notion of autonomy. As I explained in Chapter 3, neo-Kantian autonomy is narrower in scope than the Kantian notion of "autonomy in the noumenal world" and should, therefore, not be identified with positive freedom. Furthermore, the neo-Kantian philosopher's move from metaphysics (i.e. the notion of autonomy) to normative principles is not unbridgeable precisely because it is not identical with Kantian autonomy. As Richards puts it, neo-Kantian autonomy can be thought to comprise a number of "complex capacities" such as "language and self-consciousness, memory, logical relations, empirical reasoning about beliefs and their validity (intelligence), and the capacity to use normative principles in terms of which plans of action can be assessed, including principles of rational choice in terms of which ends may be more effectively and coherently realized."<sup>12</sup> Surely, the derivation of a particular social arrangement from this list of essential attributes is far more plausible than the derivation of normative principles from the Kantian noumenal world.

In my view, Galston confuses the neo-Kantian essentialist account with the Kantian noumenal world which he seems to regard as an explanationist position (i.e. as the understanding of a particular, i.e. the Western State, sub specie aeternitatis). As a result, he concludes, like

Oakeshott does, that such a position cannot yield normative principles.<sup>h</sup>

Now, even if the transition from neo-Kantian autonomy to negative freedom is valid, it seems to me that there is still a flaw in the neo-Kantian neutralist argument further down the ladder of specification: although the market economy is convenient for the neo-Kantian as it helps him to model the notion of responsibility implicit in autonomy and to reach conclusions which are not radically different from the already established institutional structure of Western democratic societies, it does not derive from the neo-Kantian essentialist account. Instead, it is injected into the argument in the late stages of specification. Perhaps the anti-neutralist claim that the market economy is really non-neutral (see 6.1) can be taken to mean exactly this; that it cannot be derived from the essences on which the higher-order, neutral morality is founded without the belated introduction of external elements.<sup>i</sup>

The role of the market in neo-Kantian foundationalism is critically examined by Raz who asserts that

[t]he individualistic bias that Rawls is accused

---

<sup>h</sup>Not surprisingly, Oakeshott himself regards the Kantian noumenal world as the product of explanationist understanding and Kant himself as an explanationist philosopher.

<sup>i</sup>It is interesting to note that the market economy can more plausibly be derived from the welfarist's essentialist account and perhaps this is why it has never been regarded as non-neutral when examined within the context of the welfarist position.

of...is not that he rules out [non-individualistic] conceptions but that he is not neutral regarding them because he makes their successful pursuit more difficult than that of individualistic conceptions of the good...The point is valid...If some market mechanism...is assumed, the value of primary goods is the function of supply and demand where the demand is partly determined by the usefulness of the goods in the implementation of conceptions of the good which are actually pursued in that society and by the number of those pursuing different conceptions. Relative to any such evaluation of primary goods some conceptions of the good will be harder to implement, i.e., will require primary goods of greater value to realize, than others.<sup>13</sup>

Raz's main point is that although the market discriminates against non-individualist conceptions by treating them as "expensive preferences", it is indeed neutral. This neutrality, however, is less than the neo-Kantian foundationalist position itself requires. It is a "narrow neutrality" in a "comprehensive conflict...which is about the ability of people to choose and successfully pursue conceptions of the good (and these include ideals of the good society or world)".<sup>14</sup> How can the foundationalist philosopher allow that market forces should interfere with the process of autonomous choice-making?

The crux of Raz's argument is that a more comprehensive neutrality is required, a neutrality which "...consists in helping or hindering the parties [i.e. the adherents of different conceptions of the good] in equal degree in all matters relevant to the conflict between them...and not just in those activities and regarding those resources that they would wish neither to engage in nor to acquire but for the conflict."<sup>15</sup> This is a neutrality of outcome which as we have

seen, and as Raz himself acknowledges, is impossible to implement.<sup>j</sup>

Clearly Raz's argument goes too far. He is not only suggesting that the market economy cannot be derived from the neo-Kantian essentialist account, but that what does derive from it is a neutrality of outcome. As I explained in 6.1, where I first examined this argument, this view is mistaken. According to the neo-Kantian position, the objects of autonomous choices do not have the same status as philosophical truth (knowledge of the essences). The idea of bringing about a situation in which the outcome of the conflict between the differing conceptions of the good is in no way determined by the operation of the basic, social institutions or by governmental policy stems from an essentialist account which is different from the one proffered by the neo-Kantian foundationalist. According to the latter, the just distribution of primary goods is compatible with all conflicting conceptions. Furthermore, a conception of the good is not thought to entail a particular list of inflexible rules and regulations or to be inextricably linked with a particular group or organization. Instead, it is revised and modified through individuals' choices. It follows that neither the resolution of the conflict between differing conceptions of the good nor their preservation in the form of isolated, closed communities is required by the foundationalist higher-order morality. What is, instead, required is the distribution

---

<sup>j</sup>This is why Raz's argument may also be seen as part of the first line of anti-neutralist criticism. (See 7.1).

of liberty rights and resources which are necessary for choosing autonomously between differing conceptions of the good. The partisans' activities aimed at attracting individuals to their particular conceptions must not violate these rights or interfere with the just distribution of resources intended to secure the (near) equal worth of liberty.<sup>k</sup>

The second type of criticism levelled against foundationalist neutrality is that its essentialist account is inaccurate. Arguments of this type do not question the meta-theoretical assumptions of foundationalism. They take for granted that there is a universal, immutable notion of human nature and challenge the neo-Kantian account of it. For instance, Raz suggests the neo-Kantian overlooks the fact that individuals have the essential second-order desire not to have their false desires satisfied.<sup>l</sup> Given that individuals choose their conceptions of the good on the basis of their desires, it follows that they do not themselves expect the state to be neutral towards their chosen conceptions. Starting from this essentialist assumption, Raz is bound to reach non-neutral conclusions.

Any essentialist account that differs from the one offered by the neo-Kantian foundationalist is bound to lead to a different set of prescriptions. For instance, Robert George

---

<sup>k</sup>See Chapter 3 on the distinction between "equal liberty" and the "equal worth of liberty".

<sup>l</sup>This point is similar to the one made by Robert George (see Chapter 4), according to which "desires are not reasons for acting".

claims that "by grounding civil liberties in true human values whose realization those liberties help to protect, promote and even make possible," we arrive at a "pluralistic perfectionism", a position which helps us distinguish sharply between morally good and morally bad conceptions.<sup>16</sup>

A similar view is put forward by John Finnis, who, like George, is an ardent critic of state neutrality. (See 4.2). According to Finnis, the "true human values" are knowledge, friendship, practical reasonableness, aesthetic experience and religion. It is from these essentially human values that he derives his index of civil rights and on the basis of which he finds the choice of certain conceptions immoral.<sup>17</sup> Such an elaborate essentialist account is bound to warrant a discriminatory conception of justice, one that favours particular conceptions of the good at the expense of others.

Let us now turn to the criticisms directed against interpretivist state neutrality. First, there are those that are levelled against **direct** interpretivist arguments (or against the particular interpretivist findings that are to be combined into a constructivist argument) and, second, those which question the coherence of **constructivist** interpretivist arguments (suggesting that the interaction between particular interpretivist findings does not really yield the principle of state neutrality).

An example of the former type of criticism is put forward by Galston who claims that Rawls' direct argument "...misrepresents what is in fact our shared cultural

understanding of personality..."<sup>18</sup> and argues that members of Western democratic societies share the view that the public understanding of moral personality must be reconciled with the non-public personality of private aims and attachments.<sup>19</sup> This amounts to saying that when Rawls talks about the distinction between the political conception of justice and the non-political conceptions of the good as being commonly acknowledged among the members of Western democracies, he mistakes what is commonly understood as an aberration for a widely shared normative principle.

A criticism of interpretivist neutrality of the second type, i.e. an argument that questions the way in which the interpretivist findings are put together into a constructivist position could be directed against the Rawlsian constructivist position. (See 5.2). But having been criticized from every other possible angle Rawls has been spared from this kind of criticism.<sup>m</sup>

The conclusion that should be drawn from this discussion is that anti-neutralism, like state neutrality itself, is not a single, coherent position. Directing against neutrality as many disparate arguments as possible does not make for a coherent philosophical position. And yet anti-neutralists have

---

<sup>m</sup>In fact the only such argument that can be found in the literature is Barry's recent criticism of Larmore's neutralist position. Barry asserts that the desire for civil peace, and the feelings of sympathy and mutual respect held together do not lead to the adoption of the principle of neutrality (Justice as Impartiality, pp. 173-177). However, as I argued in 6.3, whether Larmore's argument is interpretivist, whether, that is, he believes these three elements to be prevalent in democratic societies, is debatable.

been doing exactly that. Galston, who focuses his attack on Rawls as the main exponent of state neutrality, criticizes him for being a sceptic,<sup>20</sup> for being a foundationalist,<sup>21</sup> for being an interpretivist and, therefore, a relativist,<sup>22</sup> for putting forward the wrong essentialist account,<sup>23</sup> for giving a wrong interpretation of American public culture,<sup>24</sup> for putting forward an unacceptable position (a position that is certain to be rejected by large minorities),<sup>25</sup> and for seeking to achieve agreement rather than explore the implications of philosophical truth.<sup>26</sup> Galston cannot possibly be thinking that his is a critique of a single, coherent philosophical position. For him state neutrality is no longer a philosophical position; it is an enemy side that is prepared to use all kinds of arguments to promote the idea of a neutral state. This proves that the neutrality debate has to a great extent ceased to be a philosophical discourse and has been thoroughly politicized. And, needless to say, putting meta-theoretical constraints on political arguments is a rather foolish idea.

#### Notes

1. John Danley, "Liberalism, Aboriginal Rights and Cultural Minorities", Philosophy and Public Affairs, 20 (1991), p. 180.
2. Bruce Ackerman, "Neutralities", in R.B. Douglass et. al. (eds), Liberalism and the Good (New York: Routledge, 1990), pp. 38-9.
3. Susan Mendus, Toleration and the Limits of Liberalism. (Hong Kong: Macmillan, 1989), p. 130.
4. Brian Barry, Justice as Impartiality (Oxford University Press: Oxford, 1990), pp. 143-145; John Rawls, A Theory of Justice, pp. 328, 331.

5. Brian Barry, Justice as Impartiality, p. 144.
6. Ibid., pp. 131-132.
7. Susan Mendus, supra note 3, p. 108.
8. William Galston, Liberal Purposes (Cambridge University Press: Cambridge, 1991), p. 95.
9. Brian Barry, supra note 5, p. 169.
10. William Galston, supra note 8, p. 87.
11. Ibid., pp. 83-84.
12. David Richards, "Rights and Autonomy", Ethics, 92 (1981), p. 7.
13. Joseph Raz, The Morality of Freedom (Clarendon: Oxford, 1989), p. 119.
14. Ibid., p. 123.
15. Ibid. p. 122.
16. Robert George, Making Men Moral (Clarendon: Oxford, 1993), pp. 228-229.
17. John Finnis, Natural Law and Natural Rights (Clarendon: Oxford, 1980).
18. William Galston, supra note 8, p. 123.
19. Ibid., p. 149.
20. Ibid., p. 161.
21. Ibid., p. 127.
22. Ibid., pp. 156-157.
23. Ibid., p. 127.
24. Ibid., pp. 123, 149, 160-162.
25. Ibid., p. 130.
26. Ibid., p. 155.

## **Part IV: The Non-Normative Position**

## CHAPTER 8

### Neutrality and the Oakeshottian Civitas

The kind of state constraint position whose authorship, as I will now argue, belongs to Michael Oakeshott is remote from the battleground of the on-going Neutrality Debate. The reason for the "remoteness" of this position is its meta-theoretical background and, more specifically, the fact that it is non-normative, i.e. that it does not set moral/political principles or prescribe the establishment of a particular institutional arrangement. My starting-point will, therefore, be a brief review of the Oakeshottian view of political philosophizing, a view which I have already referred to as **explanationism**.

#### 8.1 Philosophizing About Philosophy and Philosophizing About Politics

In Oakeshott's view philosophy starts with the exploration of the conditionality of those abstractions, or "ideal characters", which have been invented to provide fixed points of reference, unquestionable criteria--the "embraces of abstraction"--facilitating our everyday lives. To question the conditionality of an ideal character is to elucidate and examine its postulates, or essences, and in their light explain why it has a certain character. A further

investigation into the postulates of the original postulates leads to an unconditional and definitive understanding, to the creation of a self-sufficient world of ideas, an absolute coherence, a concrete whole that is implied in all experiences. In Experience and Its Modes "philosophy" (along with "history", "science" and "poetry") is itself an ideal character to be understood "philosophically" whereas in On Human Conduct the ideal character is "the Western state".

My main assumption here is that for Oakeshott **political philosophy** is the understanding of "the Western state" as a complete and self-sufficient world of ideas. This involves identifying and examining its postulates. His view of political philosophy is, therefore, essentialist, in the sense that it seeks to identify the essences, i.e. the distinguishing characteristics of a given ideal character. It is also particularist, in the sense that it seeks to understand philosophically a particular political tradition, and non-normative since, for Oakeshott, philosophical understanding is thought to start where practical, or conventional, understanding ends. This kind of investigation is bound to lead "so far aside from the ways of ordinary thought"<sup>1</sup> that its findings will have no bearing on the lives and ways of the members of Western societies. They will be irreducible to moral or political prescriptions.

To use the Platonic Story of the Cave, after he makes his escape the philosopher contemplates the realities of the cave away from it and never returns. Oakeshott himself uses the allegory of the cave in order to illustrate his position but

claims that his philosopher **does** return to the cave to give an interesting account of what he has found without setting principles or advising that changes should be made in the understanding of the goings-on in the cave or in its institutional arrangements. In Oakeshott's own words, the philosopher "...who now reluctantly returns to the cave from this greatest of all intellectual adventures carries with him an unconditional understanding of the world in terms of its ultimate postulates... Thus [he] returns not with something useful in his pocket,...but with a gift of inestimable value to mankind."<sup>2</sup> Oakeshott goes on to say that if the returning philosopher were to tell the denizens of the cave that "in virtue of his more profound understanding of the nature of horses, he is a more expert horseman, horse-chandler, or stable boy than they (in their ignorance) can ever hope to be, and when it becomes clear that his new learning has lost him the ability to tell one end of a horse from the other..." he would appear rather foolish. "And if, in taking part in legal proceedings...he were to brush on one side the cave-understood conditionality of 'the truth, the whole truth and nothing but the truth' and were to insist that matters should be delayed while the question What is truth? was explored, or if he were to lecture judge and jury about the postulates of justice, those concerned might be expected to become a trifle restless."<sup>3</sup> Oakeshott does not mention the fact that there are always those who mistake the philosopher's words for prescriptive statements (after all Oakeshott himself is no exception), and who take interest in his position only insofar

as it can be shown to have normative implications. For this reason I believe it is more suitable to liken the explanationist philosopher's engagement to a permanent escape. After all, his performance before an audience of inmates, i.e. the relation of his intellectual adventures, however interesting or original or entertaining it may be, has no bearing on his engagement as a philosopher. All his philosophical work, the unconditional understanding of the goings-on of the cave, is necessarily done in the barren, empty world (devoid of any identifiable goings-on) that surrounds the cave.

The fact that Oakeshott talks of the return of the philosopher does not, of course, mean that he is not conscious of the non-normative character of his philosophy, that he, somehow, leaves it open to the chained inmates to interpret the philosopher's exposition as a normative statement. He is concerned about presenting his work as a non-normative position and for this reason he avoids using as postulates terms which signify desirable states of affairs or intrinsically valuable objectives. Instead, he replaces them with the Latin equivalents which are meant to be purely explanatory. "State", therefore, becomes "civitas", "citizen" is replaced with "cive", "public concern" is turned into "respublica" and "law" is substituted for "lex".

Given Oakeshott's rather specific, explanationist conception of political philosophizing (as the questioning of the conditionality of "the Western European state") one could say that he engages in **political** philosophizing only in the

second essay of On Human Conduct, titled "On the Civil Condition". It is there that Oakeshott identifies and investigates the postulates of the European state understood as a complete, self-sufficient world of interrelated ideas. In the present chapter I will mainly focus on this particular exposition and argue that "neutrality" or "state constraint" is actually a postulate of the Oakeshottian civitas.

The idea of state neutrality was known to Oakeshott at the time of the writing of On Human Conduct and he evidently thought it deserved a mention in his discussion of the postulates of civitas. As he puts it in "On the Civil Condition", "...the so-called neutrality of civil prescriptions is a half-truth..."<sup>4</sup> A "half-truth" is, presumably, not good enough to be a postulate of civitas, but what exactly is it? Before I give my interpretation of this rather odd statement on state neutrality, I will argue against the view that Oakeshott's is a foundationalist philosophy and, at the same time, reconstruct the context within which his statement on neutrality is made.

## 8.2 Oakeshott as Foundationalist

The starting-point of "On the Civil Condition" is the view that the idea of civitas postulates a certain mode of association between personae, namely "the Civil Association" which is a "moral association", i.e. an association in terms of a practice governed by non-instrumental rules. This is distinguished from two other modes of association, "enterprise association" and "prudential association". The former is a

relation in terms of the common pursuit of a substantive satisfaction or the achievement of a common purpose. Oakeshott mentions three main characteristics of this mode of association: First, it is voluntary. The associates have the option of quitting once they no longer share the same purpose or doubt the usefulness of the practice for the achievement of the common purpose.

Second, it is not an association of equals.<sup>5</sup> To understand this point one has only to consider the case of any one enterprise association: the members of the staff of a public hospital share the same purpose, namely the restoration of all patients to their health. For that purpose, a certain hierarchy based on expertise and experience is considered a necessary arrangement. The hospital cannot be run without the assignment of specific roles, some of which are more important than others. The contribution of physicians, paramedics and attendants to the pursuit of the common purpose is not the same and it is for this reason that they cannot be considered as equals. This sense of inequity is clearly related with "responsibility" which in enterprise association measures one's contribution to the common purpose, or the others' expectations of one's contribution.

Third, enterprise association is association in terms of the management of its pursuit. Its codes and regulations are instrumental to its purpose and can, therefore, be questioned on grounds of efficacy.<sup>6</sup>

The second mode of association to be contrasted with civil association, what Oakeshott calls "prudential

association", corresponds to the Hegelian Bürgerliche Gesellschaft and is postulated of the economy: Personae "...related to one another in continuous transaction of... their multifarious, different, ever-changing individual wants and associated in terms of a manifold of practices each composed of rules and rule-like prescriptions purporting to comprise convenient procedures for conducting the transaction entailed."<sup>7</sup> Association in terms of practices aiming at the achievement or protection of social peace is also prudential. The main point Oakeshott makes in connection with this mode of association is that, unlike the other two, it cannot be understood as a self-sufficient, complete world of ideas, or to put it differently, it cannot be understood philosophically. The reason for this is that it is composed of "...considerations to be subscribed to in choosing wants to satisfy and in performing actions to satisfy them."<sup>8</sup> Such considerations belong only to the practical mode of experience.

So "prudential association" cannot be a postulate of civitas because it contravenes the autonomous, self-sufficient character which civitas has sub specie aeternitatis. To put it differently, philosophical understanding of civitas cannot be understanding of it as a prudential association because that would not be an understanding independent of the conditionality of practice. But why does Oakeshott also reject "enterprise association" as a suitable postulate for civitas? Why does he claim that philosophical understanding of the state is understanding of it as a "moral association"?

According to the foundationalist reading of Oakeshott's position, the answer to the question lies in the first essay of On Human Conduct, titled "On the Theoretical Understanding of Human Conduct". The ideal character under investigation in that essay is not "the Western state" but "human conduct". Strictly speaking Oakeshott is not engaged in political philosophizing (as he himself defines it) but has climbed further up the ladder of condition-less understanding. The postulates here are "capacity for free agency", "reflective consciousness", "deliberation", "practice" and "morality". Starting with the first one of these, Oakeshott claims that human conduct is free in the sense that it is an expression or exhibition of intelligence.<sup>9</sup> This means not that human conduct is reflective, self-conscious or rational, but that it involves understanding and that it must, ultimately, be learned. No aspect of human conduct, that is, no understood situation, response to a situation or satisfaction obtained from others' responses, escapes this minimal condition of understanding or intelligibility. An agent's understanding of a situation is not a mere "feeling" or "organic tension" but an objective conclusion, arrived at through deliberation, open to examination both to the agent himself and to others. The agent's response is neither caused by nor deducible from the understood (un)acceptability of a situation. It is a "genuine answer" which has to be chosen by the agent himself. The latter is free precisely because his response, like the situation itself, is the outcome of an intelligent engagement. Not surprisingly, according to Oakeshott, satisfaction is also

an understood condition and not merely a "natural" enjoyment or what is commonly meant by "happiness".

Now, of all the transactional relationships between individuals there are some which are "practices", i.e. durable relationships specifying "procedural" or "adverbial" conditions to be subscribed to in acting. For Oakeshott, a practice

...does not impose upon an agent demands that he shall think certain thoughts, entertain certain sentiments, or make certain substantive utterances. It comes to him as various invitations to understand, to choose and to respond. It is composed of conventions and rules of speech, a vocabulary and a syntax, and it is continuously invented by those who speak it and using it is adding to its resources. It is an instrument to be played upon, not a tune to be played. Learning to speak it is learning to enjoy and to explore a certain relationship with other agents. The requirements of justice are not obeyed or disobeyed; they are subscribed to or not subscribed to. What Bagehot called 'coarse, causal, comprehensive usage' is a condition unknown to reflective consciousness;... Further a practice cannot itself be 'performed'. Purely regularian conduct (that is, conduct whose imagined and wished for outcome is solely that a procedure shall have been observed) is impossible: to make a grammatically faultless utterance is to say **something**...<sup>10</sup>

According to Oakeshott, there are two kinds of practices: "prudential practices" and "moral practices". The former are devices "for promoting the satisfaction of the wants of individual agents", whereas the latter are only instruments of understanding and media of intercourse. Quite obviously these two kinds of practices correspond to two of the three modes of association discussed in "On the Civil Condition": "prudential association" and "civil association". Much of what Oakeshott

says in connection with moral practices (or "moralities") in the first essay of On Human Conduct is restated in connection with one such practice, namely civitas, in the second essay of the book. The connection, therefore, between his political philosophy and his philosophical understanding of "human conduct" is the concept of "the moral practice" and of "association in terms of a moral practice". Of course, the two projects are not quite distinct from one another but the way they are related is not quite clear. My view is that if the philosophical understanding of "human conduct" is used in order to provide an answer to the question about whether the Western state is a moral or an enterprise association, then it is used as a foundation and is, therefore, at odds with Oakeshott's meta-theory.

As has been mentioned, Oakeshott holds that philosophy questions the conditionality of the various abstractions invented to facilitate the lives of those who understand the world in the practical mode. Questioning the conditionality of these abstractions means identifying and exploring their postulates or essences. The postulates can of course be understood in terms of their own postulates and the ascent of the ladder of philosophical investigation leads the philosopher to the highest metaphysical stage where, as Bikhu Parekh puts it, "[he] is only concerned with the general and permanent features of the totality of experience, and specific goings-on are no longer visible to him."<sup>11</sup> This means that political philosophy (the understanding of the Western state in terms of its postulates) is one particular stage of

philosophical understanding. It is "an inherently precarious and fragile form of inquiry. The more philosophical it becomes the further it moves away from politics."<sup>12</sup> If, in other words, the philosopher keeps questioning the postulates of civitas, he ceases to be engaged in political philosophizing and is carried "far out of sight" of civil society. Oakeshott himself is fully conscious of this point and in the first essay of On Human Conduct issues a warning to the political philosopher that he must forswear metaphysics and accept the "intermediate" nature of his enterprise.<sup>13</sup> This enterprise is the discovery and exploration of particularist, as opposed to universal, essences which explain the Western state as a homogeneous, self-contained world of ideas. Any conclusions that may be drawn from this enterprise cannot, of course, be prescriptive because none of the actual Western states is or can ever be autonomous and self-contained (i.e. independent of the economy or the pursuit of common purposes).

Now one could argue that the fact that in the first essay of On Human Conduct Oakeshott moves away from political philosophy in pursuit of universal essences causes him to lose sight of civitas. But this is not the accusation levelled at Oakeshott. The problem is not that he overshoots civitas, but that his philosophical findings are used to provide the answer to the question of whether "enterprise association" is a postulate of civitas. The point, therefore, is not that Oakeshott's "more philosophical" enterprise obscures his philosophical understanding of the Western state, but that he uses his higher philosophy as the foundation for arbitrating

between two rival philosophical understandings of civitas. And once this descent from philosophy to political philosophy is allowed, there is no reason why a further descent to the practical mode and to the formulation of moral/political prescriptions is not permissible. The implication is that Oakeshott fails to deliver a truly explanationist political philosophy.

The foundationalist reading and the ensuing critique of Oakeshott's position should not be taken lightly. The use of foundationalist language, both by Oakeshott himself and by his exponents, gives them some credibility. Oakeshott states that "civil association" is the proper postulate of civitas because "there is nothing in [it] to threaten the link between belief and conduct which constitutes 'free' agency."<sup>14</sup> In the same vein, Paul Franco, a self-proclaimed Oakeshottian, avers that civitas can only be understood as civil association because "to make enterprise association compulsory would be to deprive an agent of 'that freedom' or 'autonomy' which is inherent in agency... A state is a compulsory association. As such it cannot be an enterprise association without compromising the moral autonomy of its members."<sup>15</sup> Thus stated, Oakeshott's argument is a quintessential foundationalist argument, identical in its structure with the arguments of Chapter 3. It is Oakeshott's foundations, his essentialist account, that tell us which conception of the state (and ultimately which institutional arrangement) is "threatening" or "compromising" and which is not.

In the light of these points the very structure of On

Human Conduct, with the philosophical understanding of "human conduct" and "human agency" preceding, rather than following, the decidedly less philosophical enterprise of understanding "the Western state" in terms of its postulates, seems to strengthen the case for a foundationalist interpretation of On Human Conduct. After all, this interpretation of the book seems to be the only one that can bind all three essays together and explain the rather idiosyncratic attack on universitas (i.e. the recorded understandings of the Western state as an "enterprise association") carried out in the last essay.

In my view, the foundationalist reading may be an interesting reflection on Oakeshott's (non-philosophical) intentions in writing On Human Conduct (I believe that he tacitly encourages the foundationalist reading of his work leading to the prescription of his preferred institutional arrangement while, at the same time, maintaining the aloofness and respectability of the explanationist philosopher), but it is a misinterpretation of the philosophical argument as that appears in the book. The isolated foundationalist phrases, the order in which the essays appear and the rather Nietzschean attack on universitas in the last essay, are strong indications that Oakeshott is partial to a particular political agenda (and perhaps to a particular institutional arrangement) and that he wants his audience to know. Nevertheless, this does not interfere with or undermine his philosophical argument which remains strictly explanationist. More specifically, Oakeshott's claim that "enterprise

"association" is not a postulate of civitas does not rest on the findings of "On the Theoretical Understanding of Human Conduct". Rather, the argument is that since enterprise association is voluntary in character, in the sense that people choose to be associated in this way and given that they do not choose to be cives, i.e. that civitas is necessarily a compulsory association, the idea of a "**compulsory** enterprise association is a self-contradiction."<sup>16</sup>

Although this is the main official argument offered by Oakeshott there are two secondary ones which are also non-foundationalist in character. The first one holds that if civitas postulated enterprise association, it would not be an association of equals for its citizens would have different degrees of responsibility and, therefore, different degrees of importance and different rights and responsibilities. According to the second, if civitas had a common purpose, it would be intolerant towards individuals adhering to purposes other than its own. In Oakeshott's words, "it is not easy to rebut the view that the logic of the state thus constituted assigns to the office of its government the authority to exterminate associates whose continued existence is judged to be irredeemably prejudicial to the pursuit of its purpose."<sup>17</sup> In both of these arguments the point is not that the state qua enterprise association would compromise what is essentially human (e.g. equality or diversity), but that it is not in any way connected with the lower-order (or less philosophical) postulates of civitas, namely cives, lex, respublica etc.. It is **these** postulates that the explanationist sets out first.

Then comes "civil association" and then its postulate idea of "moral association" which is contrasted with "enterprise association" (according to the explanationist "technique" of contrasting philosophical conceptions in order to explain their homogeneous, self-contained nature). So the question of whether civitas postulates enterprise association or moral association is a pseudo-dilemma. The answer is already known since "civil association" is an already identified postulate: it is a compulsory association of equals.

Once the conceptions of "moral" and "enterprise association" have been set out, any further, more complete, philosophical understanding involves moving to higher metaphysical planes in pursuit of the series of postulates which Oakeshott introduces at the outset of On Human Conduct. Thus, we come to the ideas of practice, morality, "reflective consciousness" and "free agency" which appear to be foundations, but are actually implications of our understanding of the Western state as an autonomous, self-sufficient world of ideas (namely the lower-order, latinized postulates). Such a world, needless to say, differs from the one the citizens of Western European countries live in or can ever hope to bring about.

It is, therefore, to the cluster of lower-order postulates that I will now turn to in order to examine Oakeshott's position on state neutrality and argue that it can be seen as the missing postulate.

### 8.3 State Neutrality: The Missing Postulate?

Citizens of Western states, the Oakeshottian cives, are not enterprisers or joint enterprisers, but are related solely in terms of their common recognition of particular rules, rules which prescribe common responsibilities (and corresponding rights) and in terms of which they set aside all that differentiates them and recognize themselves as formal equals.<sup>18</sup> Setting aside those things and roles which make them different or unequal (i.e. the pursuit of particular objectives and their part in joint enterprises) does not mean renouncing or forswearing them for ever. It means that it is only as cives that they are equals and related solely in terms of rules.

These rules, the Oakeshottian lex, comprise a "system", rather than "a mere collection of rules", and relate to "the miscellaneous, unforeseeable choices and transactions" of individuals who are not joined in a common purpose, who may be complete strangers to one another, whose preferences may be as various as themselves and who may lack any moral allegiance to one another except from the one they feel qua cives.<sup>19</sup> Lex is a system of rules in the sense that it is self-contained (since it "relates those who are not, as such, otherwise related") and that there is a relationship between its prescriptions.

Of course, lex prescribes a set of abstract considerations whose relation to contingent situations must be known to and understood by cives if they are to be related as such. What must, therefore, be known and understood is

"...what will count as an adequate or acceptable subscription [to lex] in a contingent situation".<sup>20</sup> What is required in cases of uncertainty about the relation between general considerations and contingent situations is a process of adjudication, of determination of the meaning of lex, authorized in the system of lex itself.

Oakeshott makes clear that adjudication is not a bargaining process between litigants, a process whose outcome reflects the relative advantage of either one of them. More importantly, he avers that the resolution of any uncertainty, which is really an interpretation of lex, cannot be

in terms of a so-called 'social policy' or the common purpose of interest of an association: here, there is no such common purpose. The meaning of lex lies in the considerations it prescribes and in prescribing conditions lex cannot prescribe an imagined and wished-for outcome in human conduct. And if the procedure invokes a general moral consideration it must be in respect of its antecedent recognition in lex and in terms of that recognition.<sup>21</sup>

What is emphasized here is the non-instrumental character of lex which is fully congruent with the claim that individuals qua cives are equals and that they are, therefore, regarded as independent of any individual or joint purposes have set out to achieve. The terms in which cives are related (i.e. lex) cannot be determined or illuminated by the imagined or wished-for states of affairs inherent in their chosen purposes since these purposes are not set by them qua cives.

This point is reiterated throughout "On the Civil Condition". Oakeshott extends the meaning of the postulate of

lex, by adding to it the twin concepts of "civil obligation" and "civil authority", and calls it respublica, "the public concern or consideration of cives".<sup>22</sup> He identifies this more complete postulate of civitas in order to emphasize that cives are related "in terms of the recognition of rules as rules" (rather than as means to the achievement of a particular purpose or as prescriptions of rationality or wisdom).

For Oakeshott authority is attributed to respublica "not in virtue of what can exist or be achieved only in the recognition of its authority" (e.g. formality, security or social stability), nor because it is recognized to have some valuable quality or attribute (e.g. being identified with a majority of cives or with a so-called "general will").<sup>23</sup> Nor does this authority stem from the "identification of [respublica's] prescriptions with a current 'social purpose', with approved moral ideas, with a common good or general interest, or a 'justice' other than that which is inherent in respublica" or from its prescriptions' purported access to 'scientific' information about the tendencies of human action or about 'laws of historical development'.<sup>24</sup> In other words, the authority of respublica is not justified in terms of some conception of the good, be that a moral principle, a scientific theory, a foundationalist position, an independent notion of justice, a social cause or a widely acceptable social arrangement, but in its own terms. In Oakeshott's own words,

...the attribution of authority to respublica and the postulate of the evidential procedure in which the validity of its items may be determined, are nothing other than the acknowledgement of respublica as a system of moral (not instrumental) rules.<sup>25</sup>

To put it in the language of Chapter 6, Oakeshott talks of a **neutrality of justification** which, however, is non-normative in character. It is often the case that in Western societies the authority of the legal system is often justified in terms of a conception of the good (for instance, judges often claim that their ruling intends to set the right example for youngsters or to enhance family values). Oakeshott does not imply that this violates a principle implicit in social practices or that it compromises our essentially human character. Instead, he means to say that if the phenomenon that is the Western state is to be understood philosophically, i.e. in terms of its essential characteristics, the authority of its laws should not be thought to emanate from any conception of the good.

It is, of course, the case that the neutrality of civitas is also a "neutrality of aim", in the sense that, when understood philosophically the Western state does not exist in order to achieve some conception of the good or, in Oakeshott's parlance, a "substantive purpose". To understand this point it is important to examine yet another of the postulates of civitas, namely "legislation", or "the enactment of lex", and the related concept of "politics".<sup>a</sup>

---

<sup>a</sup>It should be emphasized that Oakeshott's definition of "politics" (On Human Conduct, pp. 161-163) is different from the one I offered in Chapter 2 and which I take to be

For Oakeshott "the enactment of lex" is a distinct postulate of civitas because it relates to the desirability of lex rather than to its authority. A new law is enacted and appended to respublica and an already existing one is altered or replaced, because this is regarded as desirable in a very specific sense.

First of all, the enactment of law involves an engagement which is deliberative and persuasive in character rather than demonstrative.<sup>26</sup> This means that proposed laws cannot be proven to be desirable to cives; their desirability is always questionable and has to be argued about. Arguments and actions intended to persuade that proposed laws are desirable constitute what Oakeshott calls "politics". In Oakeshott's own words, politics "...is holding up to inspection in terms of approval or disapproval some item, large or small, of respublica, considering more desirable alternatives to it, and recommending and promoting the change from what is, to the alleged more desirable condition."<sup>27</sup> Civil rulers and more specifically legislators who enact new laws do not protect or promote powerful or preferred interests but take part in politics in order to persuade others about the desirability of their proposals. And in order to be "...recognizably 'political', a proposal together with the reasons for it must relate to a possible condition of respublica and to nothing else."<sup>28</sup> This necessarily means that certain kinds of arguments are excluded from political discourse. No substantive purposes, prudential considerations, personal or

---

independent of all three views of political philosophizing.

group interest are offered as the reasons for enacting new laws or changing those already in place since all of these things are external to respublica. And to reiterate Oakeshott's main point, only respublica provides the terms in which individuals qua cives are related to each other. Reasons external to respublica are often offered for the enactment of new legislation in Western societies and in Oakeshott's view such reasons are simply non-political.

Oakeshott offers a detailed list of the reasons and proposals that are external to, or rather, incompatible with respublica. The list includes "benevolent plans for the general betterment of mankind, for diminishing the discrepancy between wants and satisfactions or for moral improvement,...proposals for awards of benefit or advantage to ascertainable individual or corporate interests claimed on account of the merit, the bargaining power or any other alleged property of such interests,...proposals to prescribe as a rule that a certain opinion, theorem, purported statement of fact, doctrine, creed, dogma, or the like to be believed to be morally right or wrong or be believed to be organically beneficial or harmful to human beings...."<sup>29</sup> It is quite clear that in this list Oakeshott includes not only what the foundationalist neutralist regards as "conceptions of the good" (including all the different foundationalisms) or contingent, institutionally irrelevant considerations, but neo-Kantian foundationalism itself which is given no special status among other foundationalist positions.

What Oakeshott also includes in the list of inadmissible

reasons for proposing new laws is what the neo-Kantian foundationalist regards as the bases for the "acceptability" or "desirability" of the neo-Kantian prescriptions (i.e. the personal or group interests within a particular society). It would be recalled that the foundationalist (like the interpretivist) thinks of the desirability of his view as irrelevant to his argument. He thinks of it as a political rather than as a philosophical consideration. In other words, it is for the politicians to popularise the neo-Kantian position, to persuade others to accept it and to establish the prescribed institutional structure. In Oakeshott's view, the kind of desirability deemed irrelevant by the foundationalist is indeed unconnected with the philosophical understanding of civitas. Nevertheless, a particular sense of "desirability" is very relevant to the explanationist view. This Oakeshott calls "civil desirability". It is desirability in terms of

a common concern that the pursuit of all purposes and the promotion of all interests, the satisfaction of all wants and the propagation of all beliefs shall be in subscription to conditions formulated in rules indifferent to the merits of any interest or to the truth or error of any belief and consequently **not itself a substantive interest or doctrine.**<sup>30</sup>

Oakeshott is keen to emphasize that the civil desirability of a proposed piece of legislation cannot be demonstrated or proven, as most interpretivists seem to suggest (see Section IV), but has to be argued about. And, to reiterate the crux of the explanationist view of the Western state, all arguments offered must be uttered in the language of respublica. The

latter could be called "neutral" precisely because it relates to a concern which "is not itself a substantive interest or doctrine", or, to use the terminology of the previous chapters, to a concern which is not a conception of the good.

For Oakeshott the Western state, so understood, could be characterised as a "mode of association" between individuals, a mode which he calls "the civil association". This is a higher-order (more philosophical) postulate leading, as has been explained, to a chain of even more philosophical concepts like "moral association", "enterprise association", "prudential association", "moral agency", "freedom", "reflective consciousness" etc.. But why does not Oakeshott ever use "neutrality" as a postulate, a concept which seems to be a suitable characterization for a mode of association which has no set purpose or direction and whose authority is not justified in terms of any conception of the good? As we have seen, this is because he thinks of state neutrality as a half-truth. It is now time to examine what he means by that.

Oakeshott's complete statement on neutrality reads as follows:

Civility...denotes an order of moral (not instrumental) considerations, and the so-called moral neutrality of civil prescriptions is a half-truth which needs to be supplemented by the recognition of civil association as itself a moral and not a prudential condition. Civil relationship is certainly a fiduciary relationship in which faithfulness is not a device for promoting the satisfaction of substantive wants; but it is not the faithfulness of friends. This does not mean that civil desirabilities are unconnected with more intimate moral relationships; it means only that what is civilly desirable cannot be inferred or otherwise derived from general moral desirabilities,

that it is not necessarily the sign of something amiss if they are not found to be pulling in the same direction or even to conflict with such desirabilities, and that political deliberation and utterance (concerned with civil desirabilities) is concerned with moral desirabilities of its own.<sup>31</sup>

Oakeshott clearly thinks that "neutrality" seems to imply that the Western state is a prudential association, that is, association satisfies the associates' multifarious wants or at the achievement of social peace. And, as was mentioned in the previous section, understanding the state as a prudential condition is understanding it in non-philosophical (i.e. in non-explanationist) terms. Unless neutrality is "supplemented by the recognition of civil association as moral and not prudential condition" it will remain a "half-truth" in the sense that it will only belong to the practical mode of understanding which can never grasp its subject (in this case the state) sub specie aeternitatis.

By saying that when state neutrality is identified with the pursuit of social peace or with the satisfaction of all individuals' wants is a-moral, Oakeshott does not mean to refute those philosophers (like Rawls) who, as we saw in the last chapter, claim that there can be no neutral (a-moral) values. What he means to say is that prudential considerations do not fit into the philosophical (explanationist) understanding of the moral, since they compromise the autonomous, self-sufficient character of the philosopher's subject.

Now, one could argue that Oakeshott's point is redundant since, as we saw in the previous chapters, state neutrality is

not connected with prudential considerations, but with a foundationalist morality, a set of interrelated moral principles stemming from what is essentially human. After all, the satisfaction of all individuals' wants (what in the previous chapter was presented under the heading of "welfarism") can itself be seen as a foundationalist position (based on the assumption that having one's wants satisfied is an essentially human experience). Oakeshott's answer to this is quite straightforward: a foundationalist position is not a "morality" because the latter, unlike the former, cannot (in the explanationist view) be deduced from universal essences<sup>b</sup> or from first principles, but is a live tradition, a language which can be learned and subscribed to by the individual in his pursuit of whatever ends. The difference, therefore, is meta-theoretical. What is initially known to the explanationist is not the essential human character perceived as a whole from which particulars can be deduced, but a tangent reality, a tradition which is to be understood as a world of interrelated ideas forming a homogeneous, self-sufficient whole.

It seems, therefore, that Oakeshott has not one but two reasons for avoiding using "neutrality": The term is used both by those who think of the state as a prudential arrangement (and this, as we saw in Chapter 6 and as Oakeshott himself makes clear, involves having a non-philosophical outlook), and

---

<sup>b</sup>One must always keep in mind that the explanationist too is an essentialist. His essences, however, are not universals but refer to the particulars which he investigates (e.g. the Western state).

by those who have a view of political philosophizing which is different from his own. And he wants to be identified with neither of these groups.<sup>c</sup> As with the substitution of the first-order postulates of civitas with their latin equivalents, he wants to make clear that the explanationist notion of state neutrality is non-normative; that it is neither political nor foundationalist.

Oakeshott's point, therefore, is that "neutrality" can indeed be regarded as a postulate of civitas, provided that the necessary clarifications are made. In this sense, state neutrality really is the missing postulate of Oakeshottian political philosophy.

Before I close this section I should make a brief comment on the suitability of "state constraint" as an Oakeshottian postulate. As I explained in Chapter 1, one of the many connotations of "neutrality" is constraint. The neutral party is constrained from acting in certain ways or pursuing particular courses of action in order to maintain its neutral status. In Part II, I used "state constraint" as an alternative to "state neutrality", an alternative which I believe creates fewer of the complications which have fuelled the neutrality debate. The question here is whether Oakeshottian neutrality as explained above connotes state constraint. It could, perhaps, be said that rulers

---

<sup>c</sup>It is of course the case that, in some sense, Oakeshott does wish to be "misread" as a foundationalist (see section 8.2). He is, however, aware that he should not achieve this by tampering with the internal coherence of his philosophy, by littering his book with terms which are used by foundationalist philosophers.

(e.g. legislators or the judiciary), when acting as such, are constrained from favouring preferred or powerful interests or from pursuing their own purposes. This, however, does not mean that the Western state can be understood as being constrained from treating its citizens in a particular way. And this is because the explanationist's state is itself the relationship between citizens and not an institutional structure whose operation affects citizens' lives. In this respect, a citizen is not someone who is treated as an essential character, but an equal member of the civil association. Drawing the line between the state and the citizen cannot but have normative implications (like the state constraint principle) and this is something that is incompatible with the explanationist task of understanding the Western state sub specie aeternitatis, that is as an autonomous, self-sufficient whole.

So far my intention has been to explain why Oakeshottian neutrality is different from the normative neutralities of the previous chapters and why it can be thought to be remote from the melee that is the neutrality debate. As I have maintained throughout this discussion, dragging Oakeshott into this debate is misunderstanding him, failing to grasp the difference between the two normative views of political philosophizing on the one end (i.e. foundationalism and interpretivism) and explanationism on the other. The final point of my argumentation will be that, in order to secure Oakeshott in his position as a truly non-normativist political philosopher, we have to make one very important distinction.

## 8.4 The "Pursuit of Intimations", Politics and Interpretivism

As we have seen, civitas postulates the idea of "politics" understood as a persuasive or argumentative engagement in which what is argued is the desirability of proposed laws. This desirability is specifically "civil" in that it is not derived from or connected with foundationalist positions, widely held views and moral principles, joint purposes or vested interests. Instead, it relates to the rules of respublica, to the public morality which provides the only terms of the association of individuals qua cives. In On Human Conduct, Oakeshott makes a negative specification of politics thus emphasizing its neutral character. According to many commentators, a positive, explanationist account of what politics is, of what it really involves, is offered in Rationalism in Politics.<sup>32</sup> It is there that Oakeshott calls politics "the pursuit of intimations" or "the pursuit of intimated sympathies". In a much cited passage he avers that "...the arrangements which constitute a society, whether they are customs or institutions or laws or diplomatic decisions, are at once coherent and incoherent, they compose a pattern and at the same time they **intimate** a sympathy for what does not fully appear. Political activity is the exploration of this sympathy; and consequently, relevant political reasoning will be the convincing exposure of a sympathy, present but not yet followed up, and the convincing demonstration that now is the appropriate moment for recognizing it."<sup>33</sup>

This view of politics sounds very familiar. It is the interpretivist view of political philosophizing, the crux of interpretivist meta-theory presented in Chapter 2. Oakeshott clearly believes that what the interpretivist understands as moral/political philosophizing is not a philosophical engagement at all. Rather, it is an altogether different concern. This can be easily explained in terms of the Platonic Allegory of the Cave. Those who engage in the pursuit of intimations are those who remain in the cave, those who are left to ponder over and argue about the meanings of the particulars of the cave while the (explanationist) philosopher is away contemplating the realities of the cave as if the cave were an autonomous, self-sufficient world on its own. In Chapter 2, I rejected this view, arguing instead, that interpretivism and explanationism are two distinct meta-theoretical views. The argument there was that the interpretivist thinks of the withdrawal from the cave as impossible, claiming that we can know of no essences, either universalist or particularist. I will now argue that although what Oakeshott calls "the pursuit of intimations" and what he does in parts of Rationalism in Politics, "Contemporary British Politics" and "On the Character of A Western State" is interpretivism, what is postulated of civitas is something that is significantly different.

My main point is that a distinction must be made between three things: First, interpretivism as a distinct meta-theory, second, "politics" as we know it and practice it in Western societies (i.e. as a language of persuasion which does not

exclude foundationalist, interpretivist or mixed arguments) and, third, "politics" as the postulate of the concept of civitas. If these three are understood as one, and I believe Oakeshott does exactly that, then explanationism loses its non-normative character. This distinction, which I suggest is essential for our understanding of Oakeshottian philosophy, is closely connected with the issue of state neutrality.

As we saw in Chapter 5, the interpretivist can engage in a large-scale investigation, in a hermeneutic project which starts with very few, if any, presuppositions. Rawls' later philosophy, exemplifies this strand of the interpretivist position. In his argument state neutrality is one of those implicit meanings, these "intimated sympathies" that he discovers and puts forward in a systematic fashion. In fact, his task is exactly this: to find those basic principles which are implicit in practice or in the existing institutions themselves. In order to make morality more coherent he looks for fundamental principles and what he finds is a distinctly political (public) morality which is not only independent of general moral outlooks concerned with the question of the good life and of personal excellence, but is thought to occupy a particular position in relation to them. Other interpretivists reject his position and find that public morality is inextricably linked with particular conceptions of the good. For Oakeshott it is not only the latter positions that are flawed, but the former ones as well. In his view, civitas postulates a sort of constrained interpretivism, an interpretivism which is necessarily limited in scope. The

Walzerian argument in favour of universal suffrage (an argument which, according to Walzer, is the epitome of interpretivism) is perfectly compatible with Oakeshott's view, but the larger scale interpretivist project undertaken by Walzer is not.<sup>4</sup> Furthermore, in Oakeshott's view, the interpretivist (or what for him is a "politician") cannot find the state non-neutral. In "Contemporary British Politics", as in "On the Character of A European State", he criticizes those interpretivists who claim that the state is an enterprise association. Although he acknowledges that it was the experience of World War II that slowly transformed the state into a centre of social planning, he refuses to accept that this realization is a valid interpretivist argument.<sup>34</sup> Clearly, this is because such arguments conflict with the explanationist understanding of the state. But this understanding is not normative in character and the imposition on "politicians" of philosophical constraints is unacceptable in Oakeshottian terms.

In "On the Character of A European State" Oakeshott comes close to realizing this. He acknowledges the validity of the

---

<sup>4</sup>To understand this point one needs only to notice the kind of "evidence" interpretivists submit with their prescriptions. Unlike the suffragettes, who only suggested that their proposed principles were implicit in existing legislation (since it treated both sexes equally), interpretivists look also for principles implicit in society's history, language and culture. More importantly, the Oakeshottian notion of "contingency" does not coincide with that of the interpretivist since it does not allow for sudden, drastic changes in the shared understandings and self-perceptions in the particular society. For the interpretivist, it is possible for the members of a democratic society to come to understand the state as an enterprise association (as in fact they did during and immediately after World War II).

interpretivist understanding of the Western state as an universitas and avers that evidence in support of the view that the Western state is a non-neutral mode of association is all around us. He even discovers a type of agent (as Rawls does in his interpretivist/constructivist argument), the "individual manqué" whom he describes as "...intolerant not only of superiority but of difference, disposed to allow in all others only a replica of himself, and united with his fellows in a revulsion of distinctness".<sup>35</sup> His dislike for this individual is, however, hardly suppressed. With strong, almost Nietzschean language he seeks to discredit him and distance himself from him. But he cannot do so as a philosopher. As an explanationist he cannot prescribe against universitas and as an interpretivist he accepts it as a sound argument. And yet he strives to find ways to oppose it. As was mentioned in section 8.3, encouraging a foundationalist reading of On Human Conduct is a way of achieving his desired effect. (According to a foundationalist argument the "individual manqué" does not have the essentially human characteristics--as the word "manqué" seems to suggest--and therefore no prescriptions can be derived from his particular characteristics).

The truth of the matter is that the interpretivist has much more freedom to investigate a particular society, a tradition, than Oakeshott suggests. And this is because interpretivism is not a postulate of civitas, but a meta-theory based on different ontological and epistemological assumptions. Unlike the Oakeshottian postulate of "politics",

it is not argumentative in character. It does not seek to persuade but to prove. Furthermore, a tradition is not presupposed since it is the core of this tradition that is to be questioned. It cannot, therefore, impose constraints on the kinds of things the interpretivist is to unearth.

Oakeshott's argument could have been coherent if he had not been tempted, like so many others, to combine his philosophy with a strong political message. Being himself so severe on those who mix politics with pleasure (that of engaging in explanationist philosophy), makes it hard to forgive him. Benjamin Barber's claim that he elevates his political agenda "...from an ideological (or psychological) preference into theoretical necessity" seems to be quite fair.<sup>36</sup>

#### Notes

1. Michael Oakeshott, Experience and Its Modes (Cambridge University Press: Cambridge, 1982), p. 356.
2. Michael Oakeshott, On Human Conduct (Clarendon: Oxford, 1990), p. 29.
3. Ibid., p. 30.
4. Michael Oakeshott, supra note 2, p. 175.
5. Ibid., p. 124.
6. Ibid., p. 126.
7. Ibid., p. 61.
8. Ibid., p. 261.
9. Ibid., pp. 32-39.
10. Ibid., p. 58.

11. Bikhu Parekh, "Review Article: The Political Philosophy of Michael Oakeshott", British Journal of Political Science, 9 (1979), p. 490.
12. Ibid., p. 490.
13. Michael Oakeshott, supra note 2, pp. 25, 33.
14. Ibid, p. 158. (Emphasis mine).
15. Paul Franco, The Political Philosophy of Michael Oakeshott (Yale University Press: New Haven, 1990), p. 190. (Emphasis mine).
16. Michael Oakeshott, supra note 2, p. 119.
17. Quoted in Bikhu Parekh, supra note 11, p. 494.
18. Michael Oakeshott, supra note 2, p. 128.
19. Ibid., p. 129.
20. Ibid., p. 130.
21. Ibid., p. 134.
22. Ibid., p. 147.
23. Ibid. p. 152.
24. Ibid., pp. 152-53.
25. Ibid., p. 153.
26. Ibid., p. 173.
27. Ibid., p. 165.
28. Ibid., p. 168.
29. Ibid., p. 168, 170.
30. Ibid., p. 172.
31. Ibid., p. 175.
32. See Paul Franco, The Political Philosophy of Michael Oakeshott (New Haven: Yale University Press, 1990), p. 193.
33. Michael Oakeshott, Rationalism in Politics (London: Mathuen, 1962), pp.123-24.
34. Michael Oakeshott, "Contemporary British Politics", The Cambridge Journal, 1 (1947-48), pp. 476-77.
35. Michael Oakeshott, supra note 2, p. 278.

36. Benjamin Barber, The Conquest of Politics (Princeton University press: Princeton, 1989), p. 171.

## Conclusion

As I first mentioned in Chapter 2, the main categorical difference between politics and philosophy is that the former seeks to persuade, whereas the latter to prove or demonstrate.<sup>4</sup> All three views of political philosophizing set out in this thesis hold that political philosophy is about revealing or discovering the philosophical truth. The foundationalist derives his normative principles from what he knows to be the truth about human nature, the interpretivist reveals principles which, he is convinced, are implicit in contemporary social practices, and the explanationist offers a superior understanding of Western political tradition viewed sub specie aeternitatis.

Each one of these views begins with different ontological/epistemological assumptions. The foundationalist

---

<sup>4</sup>It is important to reiterate the different senses of "political" that have been employed in this thesis. First, "political" is the argument that seeks to persuade and achieve the acceptability of a particular institutional arrangement. The combination of foundationalist and interpretivist elements into one position is one such argument. Second, "political" refers to principles or set of principles which refer to the operation of the institutional structure of society. Thus, Rawls talks of a specifically "political" conception of justice which, as he claims, is distinguishable from the non-political ones (those that refer to individuals' private lives) and which is prevalent in Western democratic societies. (See Chapter 5). Third, "political" refers to a particular tradition (i.e. Western democratic tradition), seen as an immutable, autonomous whole. (See Chapter 8). A "politician" is someone who puts forward "political arguments" and not someone who regards "political principles" as dictates of the foundations or as elements of endoxa.

believes that it is possible to get access to knowledge about what is essentially human, whereas the explanationist believes that what can be known is only what is essential to a particular tradition, when that is understood as an immutable, self-sufficient world, a world that does not depend on anything that is external to it. The interpretivist rejects both of these views and claims that since a moral/political tradition undergoes continuous changes, and given that these changes are contingent, any talk of essences (either universal or particular to the tradition) is fatuous. The truth can only be known at a certain point in (or, rather, in an undetermined period of) time and, therefore, the interpretivist philosopher's work never ends. Of course, this does not mean that his conclusions are tentative, but only that the truth is located in time.

Furthermore, the adherents of the three different meta-theories are also in disagreement on the question of whether philosophical truth has normative implications. The foundationalist, like the interpretivist, holds that knowledge of the truth necessarily leads to the prescription of principles regulating our lives. The interpretivist believes this to be so, because the truth for him comes in the form of normative principles (or elements of principles which beg to be put together--as in Rawlsian constructivism).

The foundationalist, on the other hand, claims that human nature can be compromised or frustrated, if certain conditions are not present. It is these conditions that he, therefore, has to set out in the form of prescriptive principles, whose

continuous specification must always be proved to be connected with the essentially human attributes.

The explanationist is the only one among the three who rejects the view that philosophical knowledge is normative in character. His claim is that since the philosopher understands the world (and more specifically politics) as something that neither is nor will ever be in an unadulterated form, it cannot be prescriptive.

A crucial point which is made explicitly in Chapter 2 is that the juggling of all three views of political philosophizing is flawed from a meta-theoretical point of view, but is common practice among politicians and philosophers with political aspirations. Even Oakeshott, whose official explanationism makes him impervious to criticism by normativists, does, at times, speak the normative idiom only to retreat to his explanationist fortress, when in fear of being involved into the inconclusive normative arguments.

What is more commonplace is the mixing of interpretivism with foundationalism with the purpose of offsetting what in politics is construed as an epistemological one-sidedness. The idea that all main principles can be derived from a minimalist account of what is essentially human, like the idea that no essences can ever be known, is hardly persuasive as a political argument, as an argument which seeks to persuade non-philosophers that one prescribed institutional arrangement (or means of bringing about such an arrangement) is better than another. The combination of foundationalism with interpretivism gives the former the ability to reach advanced

levels of specification (which would otherwise fail to reach) and the latter to avoid being identified with relativism. But to reiterate the main point, this is unacceptable from a meta-theoretical point of view.

Now, connecting this tripartite meta-theoretical distinction with the issue of state neutrality, helps to trace the philosophical arguments that are incorporated into the mixed, political arguments offered by neutralists and anti-neutralists. The foundationalist neutralist finds state neutrality to be the only arrangement that respects what is essentially human. The interpretivist, on the other hand, demonstrates that state neutrality is implicit in Western democratic societies and claims that although this may change in the future, it can be prescribed as the proper institutional arrangement. The explanationist, who like the interpretivist focuses on the Western state, avers that neutrality is an essential characteristic of the idea of the state as an autonomous, immutable world.

All three notions of state neutrality mentioned above must, of course, be distinguished from what I have called "political neutrality", that is neutrality as an argument or a policy whose ultimate goal is to make a particular proposed arrangement (foundationalist or interpretivist) acceptable to the adherents of differing conceptions of the good. Such an argument usually combines both foundationalist and interpretivist elements in order to be more persuasive. Imposing meta-theoretical constraints on such mixed arguments, by stressing that the three meta-theories cannot be combined

into one coherent position, is exhibiting political naivety. On the other hand, it emphasizes the difference between philosophy and political discourse by leading to some interesting conclusions which, however, the politician will find hard to appreciate.

The first conclusion is that state neutrality is not an amoral position, or an implication of scepticism or. Instead, it is a higher-order conception of the good, a conception that is supposed to govern the basic institutional structure of society.

Second, the endorsement of interpretivist meta-theory does not necessarily lead to the rejection of state neutrality. As we saw in 5.1, 5.2 and 6.3, the interpretivist can very well prescribe neutrality, if he finds it is predominant in the endoxa of the society he investigates. Rawls and, to some extent, Barry, Larmore and Ackerman are interpretivist neutralists. Rawls' case is, perhaps, the most significant from a meta-theoretical point of view, as he seeks to construct principles from more elementary interpretivist findings. The reasons behind this move are, I believe, political. Rawls strives to be more persuasive than the average interpretivist by proving the connection between less controversial interpretivist, finds and his principle of neutrality. But although his motives are political, Rawls' constructivist argument is really interpretivist.

Third, the neutralist philosopher need not concern himself with criticisms which find his position impracticable, unpopular or not genuinely neutral between conceptions of the

good, since these considerations are not related with the philosophical arguments themselves, but with the chances of implementing its prescriptions. The foundationalist neutralist can be criticized for having the "wrong" essentialist account as his point of departure, or for prescribing principles which cannot be traced back to his foundations (i.e. for introducing external elements during the process of specification). The foundationalist arguments discussed in Chapter 3 and Chapter 5 prescribe the market economy (a prescription which is presented as a corollary of state neutrality) without linking it sufficiently with their essentialist assumptions.<sup>b</sup>

The interpretivist neutralist, on the other hand, can be criticized for having mistaken the significant elements of endoxa and, thus, for having prescribed the "wrong" principles. The constructivist interpretivist, in particular, can be criticized for drawing the wrong conclusions from his construction of interpretivist findings. Furthermore, the interpretivist can be criticised for treating the self-interested opinions of actual members of the society he investigates as interpretivist findings from which he can derive principles to govern public life.

In Chapter 1, I argued that "neutrality" is a semantically overloaded term, and one which not better suited to the politician's purposes rather than to the philosopher's. Throughout Part II of this thesis, I replaced it with

---

<sup>b</sup>An exception to this is the welfarist position as this is presented by Barry and Raz. There, the essence of being human is the experience of having one's wants satisfied and the market economy is thought of as a way of spreading this satisfaction around in the best possible way.

"constraint" (an idea which is implicit in a certain sense of "neutrality"), in order to show that "neutrality" is a mere stage prop which the philosopher does not need to employ. There are reasons to believe that most neutralists are of the same opinion. Besides Rawls, who says so quite explicitly, most neutralists treat "state neutrality" as a technical term, whose meaning they feel compelled to clarify. We, thus, have terms like "neutrality of justification", "neutrality of outcome", "neutrality of opportunity", "neutrality of aim" etc.. (See 6.1). In my view, if the philosopher fails to make such clarifications, or to discard "neutrality" altogether, he runs the risk of being drawn into a political debate. However, in view of the fact that the Neutrality Debate is indeed highly politicized and that many a political philosopher have willingly joined in, this point betrays a certain political naivete.

My last point will be on explanationist neutrality which, being non-normative, cannot be substituted with "state constraint". Oakeshott is not impervious to criticism just because he does not prescribe moral/political principles. To the extent that he contradicts his own meta-theory (by using foundationalist language or by trying to impose constraints on interpretivist philosophy) his argument is flawed. His conception of state neutrality is, however, presented with great care in On Human Conduct so as not to be mistaken for a normative principle. For this reason it can be regarded as a truly explanationist conception. Of course this does not mean that it cannot be drawn into a political discourse by those

who speak only the language of persuasion. As for those who, like myself, find the notion of state neutrality intuitively appealing or interesting, but decide that for some reason (any of the reasons mentioned above) it cannot take the form of a normative principle, they may find it easier to accept it as the main feature of civitas, of the state understood as a world that has never and will never exist in an unadulterated form.

## Bibliography

Ackerman, B., Social Justice in The Liberal State, New Haven, Yale University Press, 1980.

Ackerman, B., "Why Dialogue?", The Journal of Philosophy, 88, 1, 1989, pp. 55-22.

Ackerman, B., "Neutralities", in Liberalism and the Good, R.B. Douglass, G.M. Mara & H.S. Richardson (eds.), New York, Routledge, 1990, pp. 29-43.

Alexander, L. and Schwarzschild, M., "Liberalism, Neutrality and Equality of Welfare v. Equality of Resources", Philosophy and Public Affairs, 16, 1, 1987, pp. 85-110.

Archard, D., "Freedom Not to Be Free: The Case of the Slavery Contract in J.S. Mill's On Liberty", The Philosophical Quarterly, 40, 160, 1990, pp. 453-465.

Aristotle, The Nichomachean Ethics, trans. J. Thomson and H. Tredinnick, Harmondsworth, Penguin, 1983.

Arneson, R. J., "Mill versus Paternalism", Ethics, 90, 4, 1980, pp. 470-489.

Arneson, R. J., "Neutrality and Utility", Canadian Journal of Philosophy, 20, 2, 1990, pp. 215-240.

Arneson, R. J., "Liberalism, Distributive Subjectivism and Equal Opportunity for Welfare", Philosophy and Public Affairs, 19, 2, 1990, pp. 158-194.

Barber, B., The Conquest of Politics, Princeton, Princeton University Press, 1989.

Barry, B., Theories of Justice, London, Harvester & Wheatsheaf, 1989.

Barry, B., "Chance, Choice and Justice", expanded text of a public lecture given at the LSE, 25 April, 1990.

Barry, B., Political Argument, London, Harvester & Wheatsheaf, 1990.

Barry, B., "How Not To Defend Liberal Institutions", in Liberalism and the Good, R.B. Douglass, G.M. Mara & H.S. Richardson (eds.), New York, Routledge, 1990, pp. 44-58.

Barry, B., Justice as Impartiality, Oxford, Oxford University Press.

Berlin, I., Four Essays on Liberty, London, Oxford University Press, 1969.

Berry, C., Human Nature, London, MacMillan, 1986.

Buchanan, A. E., "Assessing the Communitarian Critique of Liberalism", Ethics, 99, 4, 1989, pp. 852-882.

Chapman, J. W., "Toward a General Theory of Human Nature and Dynamics", in Human Nature in Politics, J. Pennock and J. Chapman (eds.), New York, New York University Press, pp. 297-319.

Cohen, G. A., "On the Curreny of Egalitarian Justice", Ethics, 99, 4, 1989, pp. 906-944.

Cohen-Almagor, R., The Boundaries of Liberalism and Tolerance, Gainsville, University Press of Florida, 1994.

Connolly, W., "Identity and Difference in Liberalism", in Liberalism and the Good, R.B. Douglass, G.M. Mara and H.S. Richardson (eds.), New York, Routledge, 1990, pp. 59-86.

Danley, J.R., "Liberalism, Aboriginal Rights and cultural Minorities", Philosophy and Public Affairs, 20, 2, 1991, pp. 168-185.

De Marneffe, P., "Liberalism, Liberty and Neutrality", Philosophy and Public Affairs, 19, 3, 1990, pp. 253-274

Dworkin, G., "Paternalism", in Philosophy of Law, J. Feinberg and H. Gross (eds.), Belmont, Wadsworth, 1980, pp. 223-241.

Dworkin, G., "Paternalism: Some Second Thoughts", in Paternalism, R. Sartorius (ed.), Minneapolis, University of Minnesota, 1983, 107-121.

Dworkin, R., A Matter of Principle, Cambridge, Cambridge University Press, 1985.

Dworkin, R., "Foundations of Liberal Equality", in The Tanner Lectures on Human Values, XI, G. B. Peterson (ed.), Salt Lake City, University of Utah Press, 1990, pp. 3-119.

Finnis, J. M., Natural Law and Natural Rights, Oxford, Clarendon Press, 1980.

Finnis, J. M., "Legal Enforcement of 'Duties to Oneself': Kant Versus Neo-Kantians", Columbia Law Review, 87, 1987, pp. 433-456.

Franco, P., The Political Philosophy of Michael Oakeshott, New Haven, Yale University Press, 1990.

Franco, P., "Michael Oakeshott As Liberal Theorist", Political Theory, 18, 3, 1990, pp. 411-436.

Frankfurt, H.G., "Freedom of the Will and the Concept of the Person", in Free Will, G. Watson (ed.), Oxford, Oxford University Press, 1982, pp. 81-95.

Galston, W.A., Liberal Purposes, Cambridge, Cambridge University Press, 1991.

George, R.P., Making Men Moral, Oxford, Clarendon Press, 1993.

Goodin, R. E. and Reeve A., "Liberalism and Neutrality" in Liberal Neutrality, R. Goodin & A. Reeve (eds.), London, Routledge, 1989, pp. 1-8.

Guest, S., Ronald Dworkin, Edinburgh, Edinburgh University Press, 1992.

Gutman, A., Democratic Education, Princeton, Princeton University Press, 1987.

Hadas, M., (ed.), Essential Works of Stoicism, New York, Bantam, 1961.

Harris, J.F., Against Relativism: A Philosophical Defense of Method, La Salle, Open Court Press, 1992.

Hobbes, T., Leviathan, Cambridge, Cambridge University Press, 1902.

Hollis, M., "The Limits of Irrationality", in Rationality, B. Wilson (ed.), New York, Harper Row, 1970, pp. 214-220.

Hollis, M., Models of Man, Cambridge, Cambridge University Press, 1977.

Hume, D., A Treatise of Human Nature, Oxford, Clarendon Press, 1988.

Jones, P., "The ideal of the Neutral State" in Liberal Neutrality, R. Goodin & A. Reeve (eds.), London, Routledge, 1989, pp. 9-38.

Kleinberg, S., Politics and Philosophy, Oxford, Blackwell, 1990.

Kukathas C. and Pettit, P., Rawls: A Theory of Justice and Its Critics, Cambridge, Polity Press, 1989.

Kymlicka, W., "Liberal Individualism and Liberal Neutrality", Ethics, 99, 4, 1989, pp. 883-905.

Kymlicka, W., Liberalism, Community and Culture, Oxford, Clarendon Press, 1991.

Larmore, C., Patterns of Moral Complexity, Cambridge, Cambridge University Press, 1989.

Larmore, C., "Political Liberalism", Political Theory, 18, 3, 1990, pp. 339-360.

MacIntyre, A., After Virtue, London, Duckworth, 1981.

Mapel, D. R., "Civil Association and the Idea of Contingency", Political Theory, 18, 3, 1990, pp. 392-410.

Martin, R., Rawls and Rights, Lawrence, University Press of Kansas, 1985.

Mason, A. D., "Autonomy, Liberalism and State Neutrality", The Philosophical Quarterly, 40, 160, 1990, pp. 433-452.

Mendus, S., Justifying Toleration: Conceptual and Historical Perspectives, Cambridge, Cambridge University Press, 1988.

Mendus, S., Toleration and the Limits of Liberalism, London, MacMillan, 1989.

Mill, J.S., On Liberty, Harmondsworth, Penguin, 1979.

Mill, J.S., The Subjection of Women, London, Virago, 1983.

Montefiore, A., Neutrality and Impartiality, Cambridge, Cambridge University Press, 1975.

Minogue, K., "Virtue, Social Justice and Moral Identity", unpublished manuscript.

Mulhall, S. and Swift, A., Liberals and Communitarians, Oxford, Blackwell, 1992.

Nott, K., Philosophy and Human Nature, London, Hodder & Stoughton, 1970.

Nozick, R., Anarchy, State and Utopia, Oxford, Blackwell, 1990.

Oakeshott, M., Experience and its Modes, Cambridge, Cambridge University Press, 1933.

Oakeshott, M., "Contemporary British Politics", The Cambridge Journal, 1, 1948, pp. 474-490.

Oakeshott, M., Rationalism in Politics, London, Mathuen & Co. Ltd., 1962.

Oakeshott, M., On Human Conduct, Oxford, Clarendon Press, 1983.

Parekh, B., "The Political Philosophy of Michael Oakeshott", British Journal of Political Science, 9, 1979, pp. 481-506.

Parekh, B., "The Rushdie Affair: Research Agenda for Political Philosophy", in The Rights of Minority Cultures, Will Kymlicka (ed.), Oxford, Oxford University Press, 1995, pp. 303-320.

Paul, J., Reading Nozick, Oxford, Basil Blackwell, 1981.

Plant, R., Modern Political Thought, Oxford, Blackwell, 1991.

Plato, The Republic, trans. D. Lee, Harmondsworth, Penguin, 1987.

Pogge, T. W., Realizing Rawls, Ithaca, New York, Cornell University Press, 1989.

Rawls, J., A Theory of Justice, Oxford, Oxford University Press, 1992.

Rawls, J., "Kantian Constructivism in Moral Theory", Journal of Philosophy, 77, 9, 1980, pp. 515-572.

Rawls, J., "Justice as Fairness: Political Not Metaphysical", Philosophy and Public Affairs, 14, 3, 1985, pp. 223-251.

Rawls, J., "The Basic Liberties and Their Priority" in Liberty, Equality and Law, S. McMurrin (ed.), Cambridge, University of Utah Press, 1987, 1-89.

Rawls, J., "The Idea of An Overlapping Consensus", Oxford Journal of Legal Studies, 7, 1, 1987, pp. 1-25.

Rawls, J., "The Priority of Right and Ideas of the Good", Philosophy and Public Affairs, 17, 4, 1988, pp. 251-276.

Raz, J., The Morality of Freedom, Oxford, Clarendon Press, 1989.

Richards, D. A. J., Sex, Drugs, Death and the Law, Totowa, New Jersey, Rowman, 1982.

Richards, D. A. J., "Rights and Autonomy", Ethics, 92, 1, 1981, pp. 3-20.

Riley, P., "Michael Oakeshott, Philosopher of Individuality", The Review of Politics, 54, 4, 1992, pp. 649-664.

Salkever, S.G., "Loppe'd and Bound: How Liberal Theory Obscures the Goods of Liberal Practices", in Liberalism and the Good, R.B. Douglass, G.M. Mara & H.S. Richardson (eds.), New York, Routledge, 1990, pp. 167-202.

Sandel, M., Liberalism and the Limits of Justice, Cambridge, Cambridge University Press, 1982.

Taylor, C., "Interpretation and the Sciences of Man", Review of Metaphysics, 25, 1971, pp. 3-51.

Taylor, C., Sources of the Self, Cambridge, Cambridge University Press, 1990.

Ten, C.L., Mill on Liberty, Oxford, Clarendon Press, 1980.

Waldron, J., "Legislation and Moral Neutrality" in Liberal Neutrality, R. Goodin & A. Reeve (eds.), London Routledge, 1989, pp. 61-83.

Waldron, J., "Minority Cultures and the Cosmopolitan Alternative", in The Rights of Minority Cultures, Will Kymlicka (ed.), Oxford, Oxford University Press, 1995, pp. 93-119.

Walzer, M., "Philosophy and Democracy", Political Theory, 9, 3, 1981, pp. 379-400.

Walzer, M., Spheres of Justice, Oxford, Blackwell, 1989.

Walzer, M., Interpretation and Social Criticism, Cambridge, Mass, Harvard University Press, 1987.

Walzer, M., "The Communitarian Critique of Liberalism", Political Theory, 18, 1, 1990, pp. 6-22.

Walzer, M., Thick and Thin, Notre Dame, University of Notre Dame Press, 1994.

Warnke, G., Justice and Interpretation, Cambridge, Polity Press, 1992.

