

REASONABLE AGREEMENT:

A CONTRACTUALIST POLITICAL THEORY

Vittorio Bufacchi

Ph.D. Thesis

London School of Economics

Government Department

June 1994

UMI Number: U062719

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 U062719

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

7144

X211248044

THESIS ABSTRACT

Reasonable Agreement:

A Contractualist Political Theory

The thesis is a defence of contractualism in liberal political theory. My aim is to show that contractualism can play a crucial role in the political theory of liberalism if it applies to the meta-ethical level rather than the ethical level. In particular, I will argue that the contractualist concept of 'reasonable agreement' provides the foundation for a new comprehensive liberal political theory. The basic intuition behind the idea of reasonable agreement is that all principles and rules must be capable of being justified to everyone: these are principles and rules on which everyone could reach agreement, where the agreement is defined in terms of what no one could reasonably reject.

The first introductory chapter will attempt to establish that contractualism reflects the ethical core of liberalism, and that the contractualist theory of reasonable agreement gives the best account of egalitarianism. This will be followed by six chapters, divided in two parts, and a brief conclusion. Part I presents the case for contractualism from a theoretical angle, providing a

conceptual analysis of reasonable agreement. Part II examines reasonable agreement from a political angle, providing an analysis of three key questions in political liberalism.

The three chapters making up Part I deal with the theories of Rawls and Scanlon, the two major figures responsible for reviving the interest in contractualism in general, and 'reasonable agreement' in particular. Chapter 2 critically evaluates Rawls's contractualism, while Chapter 3 focuses on the moral theory of Scanlon. Chapter 4 attempts to build on the efforts of Rawls and Scanlon by further exploring and hopefully improving on their theory of reasonable agreement. I believe that the strength of reasonable agreement lies in its effort to raise contractualism from the ethical to the meta-ethical level, thus the three chapters in Part I evaluate two notions central to reasonable agreement: the idea of agreement and the concept of reasonableness.

This brings us to the second part of the thesis, where the relationship between 'reasonable agreement' and political liberalism is investigated. Political liberalism is concerned with the political concepts that form the basis of a liberal society, namely, political obligation, social justice, and neutrality. Chapters 5, 6 and 7 examine how the egalitarian proposal of reasonable agreement applies respectively to these three liberal questions.

The concluding chapter will provide a summary of the main arguments presented in the thesis.

TABLE OF CONTENTS

INTRODUCTION

1. THE CHALLENGE FACING LIBERALISM

- I. The Challenge Facing Liberalism
- II. Confronting the Challenge
- III. Scepticism and the Challenge Facing Liberalism
- IV. Four Accounts of Scepticism
- V. Scepticism and Reasonable Agreement
- VI. Conclusion

PART I - REASONABLE AGREEMENT IN THEORY

2. RAWLS'S TWO SOCIAL CONTRACTS

- I. The Anatomy of a Social Contract
- II. Kant and Hobbes on the Social Contract
- III. Rawls on Social Cooperation
- IV. Rawls on Agreement
- V. Social Contracts and Theories of Justice
- VI. Conclusion

3. CONTRACTUALISM: METAPHYSICAL NOT POLITICAL

- I. The Idea of Agreement
- II. Contractualism and Moral Theory
- III. Scanlon's Critique of Utilitarianism
- IV. Scanlon's Critique of Rawls
- V. The Limits of Scanlon's Contractualism
- VI. Following in Scanlon's Footsteps

4. A THEORY OF REASONABLE AGREEMENT

- I. Some Preliminary Remarks
- II. Leaving Reasonableness Undetermined: Two Risks
- III. Defining Reasonableness
- IV. Reasonableness vs. Rationality
- V. Rethinking Reasonableness
- VI. Conclusion

PART II - REASONABLE AGREEMENT AND POLITICAL LIBERALISM

5. POLITICAL OBLIGATION

- I. Fairness and Political Obligation
- II. Natural Duty to Justice and Political Obligation
- III. Reasonable Agreement and Political Obligation
- IV. Moral and Political Obligation

6. SOCIAL JUSTICE: A COMPATIBILIST APPROACH

- I. What is Justice?
- II. Choicism and Anti-Brute Luck
- III. The Semi-Choicist Solution
- IV. The Principle of Responsibility and Justice as Fairness
- V. Cohen's Critique of the Difference Principle
- VI. Justice and Reasonable Agreement

7. ON LIBERAL NEUTRALITY

- I. Neutrality and Impartiality
- II. The Impossibility of Neutrality
- III. Neutrality and Reasonable Agreement
- IV. Neutrality and Illiberal Conceptions of the Good
- V. Neutrality and Illiberal States
- VI. Conclusion

8. CONCLUSION

- I. A Brief Summary
- II. Reasonable Agreement and the Challenge Facing Liberalism

APPENDIX 'A'. Utilitarianism, Egalitarianism and Impartiality.

BIBLIOGRAPHY

INTRODUCTION

In a review article of Charles Beitz's Political Equality, Russell Hardin makes the following claim: "Reasonable Agreement. These words are the 'open sesame' of much of contemporary political theory" (Hardin 1991, p.667). A reader not familiar with recent developments in moral and political philosophy will find Hardin's claim both surprising and puzzling. The reaction is understandable. Considering that the discipline of political theory in the western world spans over twenty-five centuries, the recent debate on the idea of reasonable agreement may appear insufficient to justify such bold claim from Professor Hardin. Yet I believe Hardin's affirmation ought to be taken seriously: his assertion is an acknowledgement of the recent widespread interest in the concept of reasonable agreement, and the belief that the concept of reasonable agreement has the power to shake liberal political theory at its very foundation.

I share Hardin's conviction (minus his irony) that reasonable agreement is where the future of political theory lies. I should

also add that Hardin and I are not the only members of this club. Recently a growing number of influential political theorists¹ have explicitly or implicitly stated their adherence to the idea of reasonable agreement.

The aim of this thesis is to argue that the contractualist theory of reasonable agreement provides the best defence of the liberal project. By 'the liberal project' I am specifically referring to liberalism as a normative political philosophy, where political institutions are evaluated on moral grounds. It follows that saying that reasonable agreement gives the best defence of liberalism means that it gives the best moral evaluation of political institutions.

While it will take all eight chapters to show that reasonable agreement can provide the moral foundations for a new comprehensive liberal political theory, the purpose of this introduction is to set out the strategy I will adopt in order to achieve my aim.

The first step will be to locate the contractualist theory of reasonable agreement in the larger political and philosophical context of liberalism. Liberalism has hitherto been concerned with the problem of reconciling diversity with peaceful coexistence, and as we enter the 21th century, this problem defines the challenge facing liberalism. The first chapter comprises a detailed account of the challenge facing liberalism, as well as a brief account of why the theory of reasonable agreement is capable of dealing with

¹ - Barry (1989b); Beitz (1989); De Marneffe (1990); Freeman (1990 & 1991), Nagel (1991), Scanlon (1982 & 1988a).

this challenge. I will also argue that the challenge facing liberalism can only be met by a theory which embraces the concept of scepticism, and reasonable agreement is such theory.

Following this first chapter, the thesis will be divided in two parts. The three chapters making up Part I will explore the theoretical dimension of reasonable agreement. It hardly needs to be said that reasonable agreement is a member of the vast family of concepts stemming from the idea of a social contract, in particular the sort of contractualism advanced by Rawls in 1971. Thus the undisputed protagonist of Chapters 2, 3 and 4 will be the concept of contractualism.

Chapter 2 critically evaluates Rawls's contractualism. Although Rawls is the major influence in the recent revival of contractualism, I believe there is a major weakness with his contractualism; namely, Rawls fails to apply contractualism to the meta-ethical level, where fundamental questions of moral intuitions and moral judgments are dealt with. Operating at the ethical level, where the rightness and wrongness of actions and practices is established, we find that Rawls's contractualism is constituted by the forced marriage of two conflicting social contract traditions: Hobbesian and Kantian. This defect has serious repercussions in Rawls's theory, thus Rawls's original position fails to capture the moral intuitions underpinning Rawls's theory of justice, furthermore it has led some liberal philosophers to doubt the validity of contractualism.

Some of the problems with Rawls's social contract theory have

been highlighted by Thomas Scanlon in 1982 in his influential article "Contractualism and Utilitarianism". Scanlon's attempt to overcome the problems inherent in Rawls's theory represents the birth of reasonable agreement as an original interpretation of contractualism. Scanlon's contractualism is the subject of Chapter 3.

Although Scanlon is the single most important reason behind the emergence of reasonable agreement as a comprehensive political theory, and arguably the most influential ambassador of the philosophical project started by Rawls in 1971, I believe there are still some problematic knots obstructing his neo-contractualist political theory. For example, Scanlon's idea of moral motivations are too dependant on the psychological faculty of 'desires'. Furthermore, the logic of Scanlon's contractualist argument is predisposed to circularity.

I believe the source of these complications lies in Scanlon's hesitation to define the concept of reasonableness, therefore Chapter 4 will be devoted to a detailed analysis of this concept. After exploring the relationship between the concept of reasonableness and the idea of agreement, I will argue that the concept of reasonableness should not be considered an agent-relative concept which defines our moral psychology, instead reasonableness is an agent-neutral concept which has the role of illuminating our moral intuitions.

This brings us to the second part of the thesis, where the relationship between reasonable agreement and political liberalism

is explored. In fact if reasonable agreement is to be considered a comprehensive political theory, it is necessary to move beyond mere definitions of concepts, and see what this theory tells us about some key political questions.

Political liberalism is concerned with the political questions which form the basis of liberal society. In particular political liberalism deals with three political questions: political obligation, social justice, and neutrality. Chapters 5, 6 and 7 examine how the egalitarian proposal of reasonable agreement as defined in Chapter 4 translates respectively onto these three liberal questions.

Chapter 5 will show that a theory of reasonable agreement provides us with an original interpretation of political obligation. I believe that as a theory of political obligation, reasonable agreement overcomes two major hurdles: justifying an obligation towards a state which provides both presumptive and discretionary goods, and explaining why a citizen may have a duty to obey a just state while at the same time preserving a right to disobey unjust laws upheld by a just state. I believe that while these two hurdles are problematic for the theories of political obligation favoured by other liberals (George Klosko and Jeremy Waldron), they are successfully cleared by the concept of reasonable agreement.

Chapter 6 looks at the implications of reasonable agreement on theories of social justice. A theory of justice must be able to reconcile two opposing but intrinsically valid principles, namely,

the principles of responsibility and compensation. I will argue that the theory of justice that derives from the contractualist theory of reasonable agreement vindicates a criterion of distribution that reconciles these two principles. I believe that such criterion of distribution finds confirmation in Rawls's theory of justice as fairness, indeed I will attempt to show that notwithstanding Rawls's rejection of desert, there are irrefutable elements of both principles of responsibility and compensation in his famous theory of justice.

Chapter 7 analyzes the concept of neutrality. The term 'neutrality' has come under extensive criticism lately, indeed today not many liberals openly defend neutrality as a cornerstone of liberal political theory. Thus we find that Rawls (1988) explicitly avoids using this term altogether, opting instead for the idea of the 'priority of right over the good'. I believe that at the root of the controversy we find a misconception of neutrality. As a result, the critics of neutrality have succeeded in defeating a straw-man, since no advocate of neutrality will ever support the kind of consequentialist neutrality critics associate with the concept.

In this chapter I will defend a principled or substantive conception of neutrality, whereby neutrality concerns a system of rules that as a whole is justified by a set of values that no one could reasonably reject. The moral and political values which define the system of rules as a whole are derived from the contractualist idea of a reasonable agreement. Furthermore, the

conceptualization of neutrality along the lines of reasonable agreement will demonstrate that neutrality plays a fundamental role in the liberal project, namely, providing the indispensable link between political legitimacy and social justice.

The concluding chapter will summarize the main arguments of this thesis, focusing in particular on the claim that the contractualist theory of reasonable agreement is equipped to confront the challenge facing liberalism.

In writing this thesis, I have incurred many debts to a number of people. My greatest debt is to my supervisor, Prof. Brian Barry, who has patiently and extensively commented on all my attempts to come to terms with contemporary political philosophy.

Considering the vast amount of literature generated by Rawls's theory over the last two decades, writing on Rawlsian themes would be a daunting task for anyone, not least for a novice as myself. Over the last four years I was fortunate to be guided along the way by Professor Brian Barry. In many ways Prof. Barry was for me what Virgil was for Dante during the latter's exploration of the Inferno: "*il savio mio maestro*" (*Inf. VIII 86*) and "*La mia guida*" (*Purg. XXVII 19*). Unfortunately the analogy ends here, as I am no Dante of contemporary political philosophy, and this thesis is certainly no equivalent of the Divine Comedy.

I am also thankful to the Economic & Social Research Council (ESRC) and the British Academy for their financial support over two years, and to all those who read or commented on parts of my

thesis, in particular Russ Bentley, Simon Caney, John Charvet, Subbu Garimella, Percy Lehning, Margaret Moore, Mike Saward and Federico Varese. Finally, I would not have finished this work without the support of Shari Garmise, whose presence has made the difference.

1. THE CHALLENGE FACING LIBERALISM

It is becoming fashionable among some political circles to claim that there is not one liberalism but many liberalisms. This proposition has been forcefully defended recently by John Gray (1989), arguably the most engaging advocate of this theory. In what follows, Gray's assertion that there are many liberalisms will be refuted.

In 1.I., I will argue that a pivotal question has challenged the liberal mind for centuries: how to reconcile diversity with peaceful coexistence. This question defines the very spirit of liberalism, indeed it brings unity to the many different liberal theories. I will refer to this question as the challenge facing liberalism.

In 1.II and 1.III, I will argue that of all the different liberal approaches, the contractualist theory of reasonable agreement can give the most persuasive response to the challenge facing liberalism. What makes reasonable agreement different from other liberal theories is that it embraces the concept of scepticism. I will argue in fact that scepticism characterises the ethical doctrine of liberalism and therefore the challenge it faces, hence a liberal theory must necessarily assume scepticism in

order to confront the challenge facing liberalism.

In 1.IV, I will critically assess and reject three different ways in which the concept of scepticism has been disapprovingly described by liberal philosophers, arguing instead for a fourth definition of scepticism along the lines of a desire to justify or give reasons to others for one's own beliefs.

In 1.V, I will argue that this definition of scepticism is intrinsic to the contractualist theory of reasonable agreement, hence reasonable agreement is the one theory able to confront and overcome the challenge facing liberalism.

1.I. The Challenge Facing Liberalism.

In Liberalisms, John Gray (1989) argues that it is impossible to ground liberalism in a comprehensive moral theory, and therefore that it is an illusion to think of liberalism as a long and cohesive intellectual tradition: "The liberalism of Locke has little in common with that of Mill and it is an error to see the two liberalisms as moments in a continuous historical process" (Gray 1989, p.262). Since there are many liberalisms, Gray concludes that a liberal political ideology is an impossibility.

Contrary to the views expressed by Gray, in what follows I will argue that there is only one liberalism, with one coherent political morality and one historical tradition. First of all, Gray's conclusion is weaken by his conception of ideology, in fact

we find that by ideology Gray understands a set of principles by the application of which ideal or good government may be realized. This conception of ideology is, to say the least, trivial, indeed it reflects a phobia of marxism which is senseless and, in this case, out of place². I believe that Gray's idea of the search for a set of principles which ought to bring about ideal or good government is better conferred by the idea of systematic theories of justice rather than ideology³.

Leaving aside Gray's understanding of ideology, it seems to me that there is little validity in Gray's argument that there is not one liberalism but many liberalisms. Instead, I believe that all forms of liberalism, *qua* liberalism, share one historical tradition and, more importantly, one ethical core. What makes liberalism a single, coherent theory of political morality is the fact that over the centuries all liberal thinkers have been looking for an answer to the same fundamental question, namely, how to reconcile diversity with peaceful coexistence. It is this very question that has come to the forefront of the political agenda recently, indeed as the increasing trend of migration from less advantaged regions of the globe will lead to greater cultural and religious diversity,

² - For an account of (non-marxist) ideology see Roger Eatwell (1993). Eatwell argues that "An ideology must possess a certain set of attributes; in particular an ideology has an overt or implicit set of empirical and normative views about (i) human nature; (ii) the process of history; (iii) the socio-political structure" (Eatwell 1993, p.7). I find his account of ideology more useful than Gray's. See also B. Goodwin (1992), Ch.2.

³ - For an account of systematic justice, see J. Fishkin (1992), Part 1.

there is little doubt that ensuring peaceful cohabitation between different people, with diverse beliefs and conceptions of the good life, will be one of the most pressing questions confronting liberalism over the next few decades.

It is this fundamental question that has challenged liberal minds for many centuries, and that gives coherence and continuity to the liberal tradition. In order to appreciate the complexity of the challenge facing liberalism it will be necessary to briefly investigate the historical and ethical dimensions of liberalism, since the challenge facing liberalism is an ethical challenge, and the ethical doctrine of liberalism cannot be dissociated from its history.

Brian Barry (1990b) points out that modern liberal institutions developed from three historical features that defined liberal states as they emerged in the 17th and 18th centuries: religious toleration, freedom of the press and the abolition of servile civil status. Modern liberal institutions may be seen as extensions of each of these features. Thus religious toleration led to the 'harm principle', freedom of the press led to freedom of expression of all kinds, and the abolition of servile civil status led to the concept of equal citizenship rights. The reason why it is instructive to start an analysis of liberalism from the historical perspective is because this approach gives us a glimpse of the ethical core of liberalism. In fact we find that the three key modern liberal institutions mentioned by Barry have the normative function of preserving the *freedom* and *autonomy* of the

individual.

The doctrine of ethical liberalism has been defined by Norberto Bobbio as follows: "[ethical liberalism is] the doctrine which accords pride of place in the scale of moral values to the individual, and hence individual freedom in the dual sense of negative and positive freedom" (Bobbio 1987, p.106). I share Bobbio's view that the ethics of individualism is at the heart of the liberal tradition. Ethical liberalism is the precondition for the other doctrines of liberalism, such as economic and political liberalism: as an economic theory liberalism is the upholder of the market system, while as a political theory liberalism upholds a theory of the state which retains the legitimate monopoly of coercive force. The point to emphasize here is that the economic and political doctrines of liberalism assume the ethical priority of the individual. Specific economic and political conditions promote the moral development of each individual, as well as the harmonious cohabitation amongst individuals who differ in aspirations, desires, needs and values. The close correlation between the ethical, economic and political dimension of liberalism reflects a crucial aspect of Bobbio's theory, namely, the belief that ethics and politics are two sides of the same coin.

While there is nothing original in this account of the liberal tradition, it was necessary to discern the ethical core of liberalism since it provides the setting for understanding the challenge facing liberalism. Once again, I believe Bobbio has pinpointed with sufficient precision this aspect of liberalism.

According to Bobbio, although individual freedom has pride of place in the scale of moral values, one person's assertion of freedom always results in the restriction of someone else's freedom. In the words of Bobbio:

[T]he problem which liberalism is called upon to resolve as an economic and political doctrine is how to make it possible for various freedoms to coexist without encroaching on each other.

(Bobbio 1987, p.106)

This is the crucial question liberals are confronted with: how to overcome the conflict between individuals without undermining their autonomy.

So far I have argued that liberalism has an ethical core. If this assertion is correct, then the challenge facing liberalism is essentially an ethical challenge. It follows that the response to this challenge must be sought in the ethical dimension of liberalism. Although the ethical doctrine of liberalism is concerned with individual freedom, it would be a mistake to jump to the conclusion that the challenge can be overcome by focusing on the concept of freedom. Since the key to the challenge facing liberalism is the idea of "coexistence", not the concept of freedom, it follows that the answer to the challenge facing liberalism lies in the idea of egalitarianism, which is a precondition of harmonious coexistence, and not in the concept of freedom.

The above assertion can be justified by the following argument: if different individual freedoms are to coexist, without one violating the other, then it is necessary for all individuals taking part to share the same egalitarian plateau. In other words equality is a pre-condition for coexistence. A coexistence that is bearable, or at least unforced, must necessarily presuppose equal concern and respect for all individuals and their respective freedom. The ultimate test for liberal theory is therefore providing an adequate conception of equality, or to put the same idea in another way, the key to the challenge facing liberalism is in providing an adequate interpretation of the concept of equality. Equality is the basis of harmonious coexistence, and thus of freedom.

In claiming that equality precedes freedom I am echoing the conclusion reached by Tocqueville (1840) in his study of Democracy in America. In an attempt to argue that democratic nations show a more ardent and enduring love of equality than of liberty⁴, Tocqueville claims that in a democracy men are perfectly free because they are all entirely equal, and they are perfectly equal because they are entirely free. This does mean that freedom and equality are undistinguishable; Tocqueville explicitly says that the taste which men have for liberty and that which they feel for equality are, in fact, two different things. Yet in democratic ages the ruling passion of men is the love of equality:

⁴ - See A. Tocqueville (1840), Vol.II, 2nd Book, Ch.1.

Among most modern nations, and especially among all those of the continent of Europe, the taste and the idea of freedom began to exist and to be developed only at the time when social conditions were tending to equality and as a consequence of that very equality. (Tocqueville 1840, p.97; emphasis added)

While I share Tocqueville's belief that equality precedes freedom, the unqualified notion of equality *per se* is too vague to be of any use. For example it may apply to outcomes, procedures or opportunity, giving rise in each case to radically different results and being justified by radically different motivations. By egalitarianism I am simply referring to the idea that all individuals ought to be treated 'as equals'; in the words of Kymlicka, an egalitarian theory attempts to "define the social, economic, and political condition under which the members of the community are treated as equals" (Kymlicka 1990a, p.5).

Over the last few years, the idea of treating people as equals, or with equal concern and respect, has increasingly become associated with the concept of impartiality, or the belief that the well-being of all moral beings matters intrinsically. It follows that the ability to justify and promote the idea of impartiality is the key to the challenge facing liberalism.

So far I have argued that liberalism has an ethical core, but that the doctrine of ethical liberalism raises a fundamental challenge to the liberal tradition. In order to overcome this

challenge what is required is a political theory that gives a convincing interpretation of the egalitarian idea. In what follows I will argue that contractualism is the only adequate foundation for a liberal political theory, since the idea of an hypothetical agreement embraces the conditions of impartiality.

1.II. Confronting the Challenge.

For the greater part of the 20th century, it was a widely held belief that utilitarianism could provide the most convincing response to the challenge facing liberalism; its combination of an objective principle (maximizing average utility) blended with an egalitarian formula (every one counts as one, no one more than one) provided an impartial standpoint that seemed to provide a valid response to the challenge facing liberalism.

It is a well documented fact that John Rawls's A Theory of Justice was generally recognised as the single most important reason behind the recent contestation of utilitarianism, and the revival of the social contract tradition as an alternative to utilitarianism. According to Rawls, the major weakness of utilitarianism lies precisely in its egalitarian formula "every one counts as one, no one more than one".

This formula is inadequate for two main reasons. First, because not all preferences or conceptions of the good life should be considered on a par. For example, your desire to stage a Nazi

demonstration chanting racist or anti-semitic slogans, and acting in a threatening fashion, cannot be judged on the same level as another person's freedom to pursue his or hers religious or cultural inclinations without harming others⁵.

Alternatively, the utilitarian formula that every one counts as one and no one more than one is inadequate because some people, due to misfortunes which are beyond their responsibility, require greater resources in order to cope with the obstacles of everyday life. In other words some of us, for example those who are physically disabled, should count more than one. Kymlicka refers to the former criticism as a failure to exclude illegitimate preferences, and the latter as a failure to recognize special relationships⁶.

Following Rawls, a growing number of liberal philosophers have accepted that the problems that undermine the utilitarian project may be overcome by contractualism. To put it briefly, contractualism looks for equality in the conditions of impartiality characterizing an hypothetical agreement which ought to be recognized as unanimously acceptable as a basis for cooperation. In other words the egalitarian character of contractualism lies in the terms of agreement which reflect impartiality and universality, therefore justifying unanimous acceptance.

⁵ - For a more detailed discussion and evaluation of this example, see Chapter 7 "On Liberal Neutrality".

⁶ - See W. Kymlicka (1990a), pp.18-44. For a detailed account of why utilitarianism gives an inadequate account of egalitarianism, and of the emergence of a non-utilitarian conception of impartiality, see Appendix 'A'.

Many liberal political theorists today believe that contractualism can give the best interpretation of egalitarianism, and therefore the best response to the challenge facing liberalism. As Hamlin and Pettit point out, the social contract tradition has been the subject of a recent renaissance: "Contractarianism in one form or another is perhaps the dominant contemporary approach to normative political theory" (Hamlin & Pettit 1989, p.11)⁷.

Of course even if we accept that contractualism, in its general form, may give a better answer to the challenge facing liberalism compared to utilitarianism, this is still not sufficient. Contractualism is not a homogeneous approach, instead there are many different theories of contractualism, which cover the full spectrum of political ideologies⁸. In order to understand why there is today a plethora of neo-contractualist political theories, it is instructive to look at the source of neo-contractualism: Rawls's A Theory of Justice.

Rawls's idea of the social contract was embodied in the (in)famous original position⁹, the most discussed and criticised aspect of Rawls's theory of justice. In Rawls's original position each person is asked to choose principles of justice, following

⁷ - Hamlin and Pettit's edited volume The Good Polity is a collection of essays by young, upcoming scholars. It is indicative that B. Barry's review article of this edited volume bears the title "The Contractarian Generation"; see B. Barry (1990a).

⁸ - This is what Hamlin and Pettit mean when they refer to contractualism "in one form or other".

⁹ - Rawls's original position will be discussed in more detail in Chapter 2 "Rawls's Two Social Contracts".

their own rational self-interest, without knowledge of their personal situation, thus excluding knowledge of their gender, race, social class etc; in other words the original position embodies the idea of rational choice from behind a veil of ignorance. Twenty years after the publication of A Theory of Justice, a curious paradox has developed. In fact those who always rejected Rawls's egalitarian intuitions have endorsed some key elements of his original position and have become strong advocates of contractualism. Thus we find that contractualism has been endorsed in order to prescribe principles of justice which from an ethical point of view oppose everything Rawls's theory stood for. I am referring here to David Gauthier, who argues in Morals by Agreement that the only true foundation of ethics is rationality, that the agreement is a form of rational bargaining, and that there is nothing wrong, from an ethical point of view, with inequalities resulting from bargaining advantages.

On the other hand those who were sympathetic to Rawls's basic moral intuitions, such as the sense of justice and the validity of fairness, have since abandoned Rawls's original position, claiming that this is at best useless and at worse counter-intuitive. While not rejecting contractualism¹⁰, this school of thought argues that Gauthier's idea of an original agreement based on conditions of bargaining and rational motivations can be replaced with another type of agreement, namely, reasonable agreement. The term

¹⁰ - This is true generally, although there are some notable exceptions, as in the case of Ronald Dworkin and Will Kymlicka who refute contractualism *tout court*.

'reasonable agreement' is generally associated with the Harvard philosopher Thomas Scanlon (1982), who argues that an agreement is reasonable if it embodies the conditions and motivation of impartiality, where such motivation is defined as the desire to justify one's actions to others on grounds they could not reasonably reject.

The theories of Gauthier and Scanlon stand at the two opposing poles in the spectrum of neo-contractualist theories of moral and political philosophy, where the former is based on rational agreement and the latter on reasonable agreement.

When I say that contractualism is the best response to the challenge facing liberalism, I don't mean that all types of contractualism are equally valid. To put it briefly, I believe that of the two recent neo-contractualist approaches above mentioned, namely Gauthier's agreement as rational bargaining and Scanlon's reasonable agreement, only the latter fully embraces the ethical spirit of liberalism and thus is in a position to respond to the challenge it is facing. One important difference between Gauthier's and Scanlon's contractualism is that unlike the idea of agreement as rational bargaining, reasonable agreement is capable of embracing scepticism, and scepticism is an essential component of the ethical doctrine of liberalism.

1.III. Scepticism and the Challenge Facing Liberalism.

What distinguishes liberalism from other political ideologies is the tolerance of conflict, and the belief that disagreements will induce to a better state of affairs. For example, it is not a coincidence that as an economic doctrine liberalism endorses the market system. The philosophy behind the market system is that conflict (i.e. competition) will lead to efficiency and fair distribution. Whether this is in fact the case or not is irrelevant. The point is that the market system, which is historically associated with the emergence of liberalism, is imbued in conflict.

The three key modern liberal institutions emphasised by Brian Barry in his account of the historical core of liberalism (the 'harm principle', freedom of expression, and equal citizenship rights) all assume that individuals will come into conflict with one another, and that the scope of a liberal political theory is to resolve such conflicts. At the same time a liberal political theory must not suppress the conditions for individual disagreements, instead the possibility of disagreeing with one another must be preserved, since such disagreements (within certain limits) are the measure of a healthy liberal democracy.

This view is usually associate with J.S. Mill's liberalism, who warns us in On Liberty against the tyranny of the majority, and the importance of dissenting views. According to J.S. Mill, diversity of opinion is a *sine qua non* condition for the preservation of liberty and the moral development of all individuals, indeed J.S. Mill felt that we have an obligation to

challenge our beliefs, even if there was no opposition to it:

If resistance waits till life is reduced nearly to one uniform type, all deviations from that type will come to be considered impious, immoral, even monstrous and contrary to nature. Mankind speedily becomes unable to conceive diversity, when they have been for some time unaccustomed to see it. (J.S. Mill 1972, p.142; emphasis in original)

Isaiah Berlin (1969) rightly points out that at the root of J.S. Mill's argument there is an epistemological position, namely, "[T]hat human knowledge was in principle never complete, and always fallible; that there was no single, universally visible, truth" (Berlin 1969, p.188). Thus diversity of opinion is not upheld by J.S. Mill because of its instrumental value, that is to say because it takes us a step closer to the single truth, but instead diversity of opinion is desirable for its own sake¹¹. It is, in a sense, what liberal morality is all about.

Diversity of opinion is, according to J.S. Mill, a founding stone of liberal democracy. Indeed the acceptance of reasonable conflict is the essence of a liberal world. Not surprisingly we find a similar message to that of J.S. Mill in the work of Stuart Hampshire (1983):

[M]orality and conflict are inseparable: conflict between

¹¹ - See I. Berlin (1969) p. 189-190.

different admirable ways of life and between different defensible moral ideals, conflict of obligations, conflict between essential, but incompatible, interests. I believe now that the subject-matter of morality is misrepresented, and finally disappears from view, when a moralist concludes with a picture of the ideal human life and of a possible harmony of essential human interests, as Aristotle and Spinoza both did. I suggest reasons to disbelieve that there can be any such single ideal and any such ultimate harmony. (Hampshire 1983, p.1)

If we accept the above characterization of liberal political theory as imbued in conflict, it follows that the challenge facing liberalism, far from being inimical to the liberal project, is intrinsic to the liberal tradition, indeed it reflects the very spirit of liberalism: the ethical core of liberalism and the ethical challenge facing liberalism cannot be separated.

Earlier I defined this challenge in terms of how to make it possible for various freedoms to coexist without encroaching on each other. The assumption here is that individual freedoms will come into conflict, and therefore a solution to this challenge must resolve such conflicts. Yet it is important to emphasize that such conflicts must be reconciled without destroying the preconditions of conflict. In other words, without undermining the conditions of diversity, since without diversity there would be no liberalism. Another way of expressing the same point is by saying that

liberalism must overcome but tolerate reasonable disagreements.

I hope to have shown why the challenge facing liberalism (indeed the ethical spirit of liberalism) and the idea of conflict are correlated. If we accept this premise, then we must also accept the idea that liberalism must necessarily rest upon a commitment to scepticism, since scepticism is the precondition for accepting diversity. It ought to be emphasized that we cannot have impartiality unless we endorse a form of scepticism. If by impartiality we understand the belief that the well-being of all moral beings matters intrinsically, hence that all interests ought to be given equal weight, then scepticism is a necessary condition of impartiality.

Scepticism here does not refer to our judgment of the views held by others, instead it refers to the way we see our own beliefs and preferences. It is only by endorsing a sceptical frame of mind that the motivation of impartiality is fulfilled. Unless a liberal theory embraces scepticism, it is doubtful that it will be able to promote impartiality as the basic criterion of moral motivation, hence give the best response to the challenge facing liberalism. Of all recent neo-contractualist theories, only reasonable agreement is in a position to embrace the conditions of scepticism, therefore only reasonable agreement is in a position to respond to the challenge facing liberalism.

While it will take the entire thesis to show why reasonable agreement gives the most persuasive reply to the challenge facing liberalism, in the remaining pages of this chapter I want to argue

that any liberal political theory must first demonstrate that it has the quality of accepting the conditions of scepticism. As I will argue shortly, a sceptical frame of mind is a condition on which disagreement and reasonable conflict necessarily count on, hence scepticism is intrinsic to the spirit of liberalism as an ethical doctrine.

In what follows I will first analyze the unhappy relationship between the majority of contemporary liberals and the doctrine of scepticism, followed by an account of the type of scepticism which I feel liberalism ought to be committed to. Finally, I will show why reasonable agreement endorses scepticism, and its implications for the challenge facing liberalism.

1.IV. Four Accounts of Scepticism.

The concept of scepticism is one of the oldest concepts in the Western culture. Scepticism was a hot topic in ancient Greece, being explored by both Plato and Aristotle in the course of their lives. In the first century BC scepticism became recognised as a formal school of thought, indeed it came complete with a guru figure (Pyrrho of Ellis), disciples (Sextus Empiricus) and a manifesto (*Outlines of Pyrrhonism*).

Since then scepticism has been at the forefront of philosophical debates, and all the great western philosophers inevitably felt compelled to confront this issue. Descartes devoted

his life to fight it; while David Hume attempted to overcome the challenge of scepticism while acknowledging its validity; Kant felt he had to dealt with it. In our century many distinguished philosophers have been intrigued by it, thus Bertrand Russell wrote a series of essays entitled Sceptical Essays; Moore felt the need to quell scepticism; and Wittgenstein in his late works On Certainty was unsure whether scepticism could ever be defeated. This array of philosophers ought to be sufficient to ensure that scepticism remains one of the key questions in political philosophers, although it must be said that many contemporary liberals have treated scepticism with little respect.

It is not my intention to present a short history of scepticism. The point is that any concept that withstands the test of time is a powerful concept, and must be taken seriously: scepticism is such concept. The power of scepticism can be seen by two examples which reflect the immense impact the concept of scepticism had on history.

The first example argues that the political implications of scepticism lie in the direction of a radical egalitarianism. This argument has been defended by Aryeh Botwinick (1990), who claims that endorsing scepticism goes a long way towards justifying the expansion of political participation:

If skepticism in some of its different guises states that none of us is in a position to affirm the rational superiority of its views or values over those of his fellows, then the

appropriate political response is to have as many members of society as possible participate in the numerous collective decisions affecting our lives. Skepticism delegitimizes the formation of any permanent hierarchies in society and provides a continually renewing impetus for the expansion of political participation. (Botwinick 1990, p.7)

The above argument finds support in the events of the French Revolution, indeed one can say that scepticism played a crucial role in shaping the course of European history. At the end of the 18th century, scepticism was invoked by the political 'Left' in France to undermine the despotic regimes clinging to the status quo and finding legitimacy in the idea of divine right to rule. In this context endorsing scepticism meant raising doubts on the rule of the social, economic and political elites, doubts on the legitimacy of despotic monarchies, doubts on the conduct of the Church. All these doubts were transformed, with time, into challenges: first intellectual challenges and then physical challenges.

It is a long-enduring controversy whether there was a causal relationship between the philosophy of the Enlightenment and the French revolution. I don't know what the answer to this question is, yet I believe that it can be affirmed with some degree of certainty that the French revolution would not have taken place if it was not for a fair dose of scepticism. To cut a long story short, we must recognise scepticism as a progressive and revolutionary force; scepticism was the epistemological foundations

behind the fight against all kinds of dogmas, and thus behind the birth of liberal democracies.

Consider now the second example. Scepticism has been adopted by the political 'Right' in the realm of international relations to justify the status quo. In the discipline of international relations those who call themselves 'realists' appeal to moral scepticism to justify three inter-related claims: (1) the image of a Hobbesian global state of nature, where all nations are moved exclusively by self-interest; in fact it is claimed that a ruler has an obligation to follow the interest of its nation; (2) the absence of moral norms governing relations between states and, *dulcis in fundo*; (3) the claim that might makes right. 'International moral scepticism' refers to the idea that there is no room for morality in international matters. Scepticism is thus used in this second case to inflict doubt, or even denial, concerning the validity of moral arguments. Scepticism justifies the present amoral status quo in international relations¹².

These two apparently contradictory examples are an indication of the impact of scepticism on our lives. Analysing the impact of the concept of scepticism on French revolution or international relations would require another thesis, and I cannot hope to do justice to either argument here. Instead, the only conclusion I want to draw from these two examples is that scepticism is a powerful concept, which deserves to be taken seriously.

Notwithstanding the impact of scepticism on our history, it is

¹² - See Charles Beitz (1979), especially Part I.

surprising to find that many liberal thinkers have felt the necessity to distance themselves from this doctrine, denying any sort of affinity between liberalism and scepticism. In what will at first appear a hopeless attempt to swim upstream, I will argue that scepticism is one of the founding pillars of the liberal tradition. That is to say, I don't simply believe that there is no contradiction between liberalism and scepticism, but that one can only endorse liberalism to the extent that one endorses scepticism; in other words, embracing scepticism is a *sine qua non* condition for embracing liberalism¹³. In what follows I will investigate the meaning of this concept and its turbulent relationship with the liberal tradition.

It is sufficient to take a brief overlook at the literature on scepticism and liberalism to understand why scepticism has been considered by many advocates of liberalism to be the weak link in the liberal chain; namely, because scepticism has been unfairly misconceived. Liberal conceptions of scepticism can be listed under four headings: (1) scepticism as denial of truth; (2) scepticism as rejection of absolutes; (3) functional scepticism; (4) scepticism as an invitation to justify beliefs. Of the above four definitions of scepticism, I will argue that only the fourth one has any validity, the other three being straw-man formulations needed to make the target of certain critiques easy to centre.

By supplying this short list of liberal conceptions of

¹³ - Although the opposite is not always true, that is to say, one can endorse scepticism without endorsing liberalism.

scepticism I am not claiming to have exhausted all the possible shades of meaning attributed to scepticism, or that there are no other ways in which scepticism can be defined. Nevertheless I do believe the four positions I have picked out cover the greater range of possible definitions, and other definitions can be traced back to one of these four conceptions.

(1) Scepticism as Denial of Truth.

It is a widely held view that scepticism implies the negation of objectivity, indeed this view is shared by Ronald Dworkin, one of the most influential contemporary liberals philosophers.

Dworkin discusses scepticism as part of his critique of Ackerman's conception of liberal neutrality. The concept of scepticism is central to Dworkin's argument, indeed his rejection of Ackerman's notion of liberalism as neutrality rests on the fact that neutrality makes a fatal opening to scepticism, and scepticism is inimical to liberalism. In the words of Dworkin, Ackerman is wrong to ground liberalism on neutrality "because it makes liberalism much more vulnerable to the familiar charge that it rests on moral scepticism or nihilism" (Dworkin 1983a, p.47).

It is not clear if, by using the conjunction "or" between the terms "moral scepticism" and "nihilism" in the above quotation, Dworkin understands moral scepticism and nihilism to be synonymous. If this is the case, he is quite simply wrong; even though there may have been an overlap in the history of these two ideas, no

modern advocate of scepticism would call himself a nihilist, furthermore it is important to distinguish between a political doctrine (nihilism) and an epistemological doctrine (scepticism).

Leaving aside the question of scepticism and nihilism, I want to concentrate on Dworkin's account of scepticism. I believe that by scepticism Dworkin understands some extreme form of subjectivism, indeed Dworkin defines scepticism as follows: "scepticism ... argues that beliefs about how people should live are merely 'subjective' and have no 'objective' validity" (Dworkin 1983a, p.47). It follows that there are two reasons why Dworkin rejects scepticism. Firstly because he believes that scepticism may open the door to indifference; here Dworkin seems to agree with Rawls, who rejects the claim that the idea of an overlapping consensus on a political conception of justice "implies indifference or scepticism as to whether a political conception of justice is true" (Rawls 1987, p.12). Secondly, and more importantly, Dworkin is afraid that scepticism denies him the right to stand up to those who hold views which he feels are obviously wrong:

Liberalism cannot be based on skepticism. Its constitutive morality provides that human beings must be treated as equals by their government, not because there is no right and wrong in political morality, but because that is what is right.

(Dworkin 1985, p.203)¹⁴

In order to strengthen his argument, Dworkin mentions the case of the moral majority which believes that heterosexuality is right and homosexuality wrong. Dworkin feels that by endorsing scepticism, liberalism gives up the possibility of challenging the majority on the grounds that its views are wrong. While I sympathize with Dworkin's conclusions, it seems to me that his attack on scepticism is out of place. Dworkin is objecting to an historical conception of scepticism, which few sceptics would uphold anyway. Rejecting extreme forms of subjectivism does not entail rejecting scepticism *tout court*, but only its ancient (Pyrrhonist) conception.

Equating scepticism with nihilism, indifference or denial of truth is like denying scepticism any conceptual development since its Pyrrhonist golden days in the first century AD. Dworkin does not seem aware of the fact that scepticism has changed radically since the days of Pyrrho of Ellis or Sextus Empiricus, indeed no-one should be surprised by this development nor hold it against the doctrine of scepticism. After all the concept of democracy also stems from Ancient Greek roots, yet no modern-day champion of democracy would claim to be pursuing the kind of democracy advocated by our Hellenic predecessors, nor would anyone criticise

¹⁴ - Dworkin also claims that grounding liberalism on some form of moral scepticism makes liberalism vulnerable to the charge that it is a negative theory for uncommitted people, furthermore it offers no effective argument against justifications for economic inequality. See Dworkin 1983b.

modern democracy on the grounds that its Hellenic ancestor was undesirable; yet this is what Dworkin's objection to scepticism amounts to.

I will have more to say about the difference between ancient and modern scepticism later on in this chapter (definition no.4), for the time being let it suffice to say that equating scepticism with denial of objectivity or even nihilism reflects a very narrow understanding of the concept of scepticism. Scepticism is more sophisticated than Dworkin seems to believe, indeed far from undermining the doctrine of scepticism, abandoning the Pyrrhonist legacy can only be beneficial to its modern version.

(2) Scepticism as Rejection of Absolutes.

The second type of scepticism in our list takes the form of the rejection of absolutes. This interpretation is considered by H. Richardson (1990), who goes on to argue that this particular sceptical stance is distinguished by its self-defeating nature. In order to highlight the incoherent nature of the sceptical argument, Richardson considers the problem of how liberals ought to treat the militantly intolerant:

[I]f the liberal refuses to own up to such strong normative assumptions [i.e. that rule out giving full political sway to the conceptions of the good of its illiberal citizens; V.B.], and instead rests on a sceptical rejection of 'absolutes',

then on what reasonable grounds are illiberal conceptions of the good held to be false? (Richardson 1990, p.4)

Richardson's critique of scepticism is that the concept of scepticism itself is an absolute (i.e. intolerant) belief, hence if scepticism is a rejection of absolutes, paradoxically it is also a rejection of scepticism. Richardson concludes by claiming that scepticism has a self-defeating nature, since it fails to clear the most basic logical fence: coherence. A similar critique of scepticism to Richardson's can be found in Hilary Putnam (1981), who claims that extreme forms of scepticism and relativism are self-refuting suppositions, where a self-refuting supposition is one whose truth implies its own falsity:

For example, consider the thesis that all general statements are false. This is a general statement. So if it is true, then it must be false. Hence, it is false. (Putnam 1981, p.7)

My impression is that Richardson's critique of scepticism is based on a caricature of the concept. The only way for Richardson's critique of scepticism to have any validity is for scepticism to somehow imply the denial of beliefs, since by logical implications the doctrine of scepticism itself would not pass its own test. My objection to Richardson's account of scepticism echoes what I said earlier about Dworkin's critique, namely, it fails to distinguish

ancient from modern scepticism¹⁵.

Modern scepticism is a doctrine which questions the epistemological foundations of our beliefs, it does not deny the validity of our beliefs. In other words while 'denial' was the key concept in Pyrrho's doctrine, 'questioning' is the key concept in modern scepticism. It follows that the assumptions of Richardson's critique are invalid. Scepticism cannot and does not distinguish between extreme and non-extreme beliefs, or in the language of Richardson between 'absolutes' and, I assume, 'moderate' beliefs. Modern scepticism does not discriminate between beliefs, instead its only function is to question all beliefs *qua* beliefs.

(3) Functional Scepticism.

The third type of scepticism I will call functional scepticism. This version of scepticism was attributed to Bruce Ackerman by Simon Caney (1991), who rests his critique of Ackerman's defence of neutrality on the unpersuasive nature of functional scepticism. In what follows I will argue that Ackerman did not advocate functional scepticism, and that functional scepticism is not a correct reading of scepticism.

In his "Consequentialist Defences of Liberal Neutrality" Caney

¹⁵ – My critique of Richardson does not extend to Putnam. Since Putnam does not consider more moderate forms of scepticism, but only the extreme scepticism of Sextus Empiricus, it would be wrong to conclude that Putnam does not distinguish modern from ancient scepticism. There is nothing in Putnam's book to suggest that all forms of scepticism, even non-extreme scepticism, are self-refuting suppositions.

argues that, in contrast with what advocates of liberal neutrality want to believe, the neutral state does not have better consequences than a perfectionist state. In order to persuade us of his thesis, Caney examines a variety of arguments supportive of liberal neutrality, showing how each one is incoherent, unrealistic or, more simply, unappealing. One of the issues he discusses is the "importance of doubt" argument advocated by Ackerman; the basic idea behind this argument is to ground neutrality on the idea of scepticism.

According to Caney, the way Ackerman defends neutrality is on the grounds that a worthwhile life involves the availability of worthless options:

A life can only go well, he [Ackerman] claims, if individuals have the possibility of making mistakes and this requires in turn that the state does not ban unworthy forms of life.

(Caney 1991, p.468)

It follows from this argument that the importance of doubt, or scepticism, is related to the choice of worthy options:

Similarly, Ackerman argues, genuine ethical worth requires that one is open to doubt and that one therefore chooses worthy options for the right reason. This requires in turn that the state desist from banning worthless options - hence neutrality is justified. No form of the good life should be

banned on the grounds that it is unworthy because even unworthy conceptions of the good serve a role. Thus the state should be neutral between conceptions of the good. (Caney 1991, p.468)

My reading of Ackerman is different from Caney's. The importance of doubt is not that it allows people to choose a worthy option for the right reason, i.e. because they can compare it with unworthy options. Instead doubt is important because it pushes us to question our beliefs, thus developing a critical perspective. Doubt is what provokes dialectical engagement and response: "in short, rather than fearing rigorous dialectical self-examination, the liberal's last, best, hope lies in the encouragement of philosophical reflection" (Ackerman 1983, p.388). If my reading of Ackerman is correct, the importance of doubt in Ackerman's liberalism may be compared to Rawls's views on the importance of being able to revise our conception of the good¹⁶.

I labelled Caney's interpretation of Ackerman functional scepticism. By using the term 'functional' I am referring to the idea that doubt introduces unworthy options, and unworthy options have the function of providing us with the reasons for choosing worthy options. As I said, I don't think that this is what Ackerman had in mind when he was talking about doubt, but for the sake of the argument let us assume that Caney is right in his interpretation of Ackerman, that is to say, the role of 'doubt' in

¹⁶ - See also Kymlicka (1989a), pp.15-19 and 51-61.

Ackerman's liberalism is purely functional.

If this is the case, then I have no problem rejecting Ackerman's functional approach to scepticism for being quite simply wrong. Scepticism and functional explanations cannot go hand-in-hand, for the simple reason that scepticism is a epistemological doctrine, that is to say, a doctrine that questions the firmness of our beliefs, while functional explanation is, by definition, an explanation. Scepticism is a prescriptive doctrine, and as such it is not concerned with explanations. For this reason scepticism cannot have the functional role it was attributed by Caney. If scepticism has any function it is simply to induce us to question our beliefs so that we can revise and improve on them; nothing more.

(4) Scepticism as an Invitation to Justify Beliefs.

So far I have assessed three criticisms of the concept of scepticism all stemming from weak and unrealistic interpretations of the concept. I rejected the idea of scepticism as denial of truths on the grounds that it refers to an archaic (pyrrhonist) conception of scepticism. I have rejected the idea of scepticism as rejection of absolutes on the grounds that scepticism does not discriminate between beliefs on the basis of their degree of extremism, instead it questions the foundation of all beliefs qua beliefs. Finally I have rejected the idea of functional scepticism on the grounds that scepticism is an epistemological (prescriptive)

doctrine, not a form of explanation. So what is the correct reading of scepticism?

I believe that modern scepticism is a doctrine which simply challenges us to provide justification for our beliefs. The point to stress here is that, contrary to what is generally believed, scepticism is not a claim: it does not claim that we don't know anything, nor that we cannot know anything, or that truth doesn't exist. Scepticism does not provide answers, it only raises questions.

To define scepticism in terms of an invitation to justify beliefs has some notable strengths. Its best asset is that it holds a wide consensus among both its advocates and its adversaries in philosophical circles¹⁷. For example Christopher Hookway (1990) argues that the challenge for justification is the most striking distinguishing mark between ancient and modern scepticism. As I showed earlier, this distinction was overlooked by some critics:

[A]ncient scepticism differs from its modern versions. Where twentieth-century philosophers interpret scepticism as the doctrine that none of our beliefs are justified, or that none of them count as knowledge, ancient scepticism seems more

¹⁷ - To say that it holds a wide consensus does not mean that it holds unanimous approval. It will suffice to provide one example: James Franklin (1991) argues that scepticism is about the impossibility to distinguish the true from the false, and the best argument for scepticism is that from symmetry. There may well be other arguments for scepticism, although for the purpose of this chapter it is not necessary to list them all; I am only interested in justifying scepticism defined as an invitation to justify beliefs.

radical. The Pyrrhonist attacks the possibility of belief: his argument is directed towards persuading us to suspend judgment on all things. (Hookway 1990, p.2)

A.C. Grayling (1985 & 1991), who is far from being sympathetic towards scepticism, also accepts a similar definition.

But the most troublesome scepticism claims nothing: it merely challenges us to provide justifications. The sceptical arguments are designed to show why a defence of justification is needed, not to prove a negative thesis. (Grayling 1991, p.24)

Apart from being the most accurate interpretation of modern scepticism, the definition of scepticism as a challenge to justify our beliefs has the advantage of being more appropriate to the context of political philosophy rather than epistemology. After all the reason for the present discussion of scepticism is not motivated by a desire to discuss epistemological questions, instead it originates from the attempt to argue that scepticism is integral to a political concept of liberalism; it is not a coincidence that my critique of adversaries of scepticism were all political theorists.

Defining scepticism in terms of a challenge to justify one's beliefs has a second important characteristic, namely, it applies to first or primary order beliefs, not to second order beliefs.

This is the fundamental difference between the concept of scepticism, as I have defined it, and the concept of fallibilism. In a recent article, Raz (1989) criticizes the view that scepticism provides the moral foundation for respecting individual liberty. What is interesting to note is that Raz distinguishes scepticism from fallibility¹⁸: scepticism implies that no knowledge can be gained on any moral value or on certain particular issues, while fallibilism is the realization that our beliefs may be mistaken. I think this distinction is important, although my definition of scepticism is different from Raz's.

Acting on first-order beliefs, scepticism does not tell us whether our beliefs are indeed right or wrong, instead it questions the way we come to hold certain beliefs. For example, I hold the belief that whites and blacks are equal, hence they ought to have the same rights. Fallibilism would question the belief that in fact people of different races are equal, while scepticism (as I have defined it) questions the reasons why I believe that whites and blacks are equal. The advantage of distinguishing between scepticism and fallibilism is that once I justify my first-order beliefs, I do not need to worry about charges of fallibilism. This is important since it indicates that moral scepticism does not lead to moral relativism. On the contrary, our moral beliefs are strengthened by the process of being challenged to justify our beliefs to others.

¹⁸ - "The realization that our beliefs may be mistaken should not, however, be confused with skepticism" (Raz 1989, pp.763-764).

To summarize my argument so far, I have argued that the best definition of scepticism is the following: scepticism is the doctrine which challenges us to justify our beliefs. My next step is to show that this conception of scepticism is pertinent to liberal political theory, indeed that there is a strong correlation between scepticism and the idea of reasonable agreement.

1.V. Scepticism and Reasonable Agreement.

Before I proceed it may be useful to recall why I have entered into this digression on scepticism. Briefly, it is because I believe scepticism is a condition for the fostering of impartiality as a moral motivation, or in other words, scepticism is the epistemological assumption behind the ethical doctrine of liberalism. It follows that acknowledging the condition of scepticism is a prerequisite for any liberal political theory to successfully respond to the challenge facing liberalism. In what follows I will first try to spell out why I feel scepticism, defined as a challenge to justify our beliefs, plays a central role in the liberal project. This will be followed by an account of why of all neo-contractualist political theories, only reasonable agreement is capable of embracing the condition of scepticism.

Earlier I have concentrated on arguments by liberal thinkers that denounce the correlation between liberalism and scepticism. What I didn't say is that there are some liberals who swim against

the current. The liberalism of Bruce Ackerman (1980) is arguably the best known case of an attempted marriage between liberalism and scepticism, indeed Ackerman needs to endorse scepticism as part of his attempt to show that neutrality is the central concept in a liberal theory.

It seems to me that there is a problem of circularity in Ackerman's position, concerning the relationship between scepticism and the concept of neutrality. That is to say, his defense of scepticism is based on his conception of liberal neutrality, and his defense of liberal neutrality is based on his conception of scepticism. In other words, instead of deducing a theory of liberalism from the assumption of scepticism, by the time Ackerman introduces the argument for scepticism, his theory of liberal neutrality is all but complete.

Contra Ackerman, I want to argue that scepticism precedes liberalism. Among contemporary liberals, I believe Peter Jones (1989a) comes closest to endorsing a similar position, indeed Jones argues that reasonable doubt may have a significant role to play in liberal argument. Jones sees freedom of belief as the central feature of the liberal argument, and he argues that each individual has an interest in having his or her beliefs respected simply because those are his or her own beliefs. According to Jones, this concept of freedom of belief can be found in the deontological liberalism of Rawls and Dworkin, as well as in the perfectionist liberalism of Haksar and Raz, although these two types of liberalism are more plausible against a background of reasonable

doubt and reasonable disagreement.

While I agree with Jones's general line of argument, my views differ from his to the extent that I am taking a more extremist position. Although Jones is making an argument for reasonable doubt and liberalism, he is also very careful to emphasize that he is not arguing that liberalism necessarily entails scepticism, or that scepticism must issue in liberalism. In the closing paragraph of his paper Jones writes:

However, I have not argued that the claim that there is scope for reasonable doubt and reasonable disagreement must, of itself, translate into an assertion of liberalism. Much less have I argued that liberalism either requires or must itself issue from a more general and more profound scepticism. (Jones 1989a, p.69)

I don't see what Jones is gaining by making such disclaimers on his theory. With the risk of drawing heavy criticisms, my approach to the question of scepticism and liberalism will be more extreme. I will attempt to argue that, to the extent that it fosters diversity and emphasizes the reality of the challenge facing liberalism, scepticism is a founding pillar of the liberal project.

I have argued before that the toleration of conflict and disagreements is what distinguishes liberalism from other political ideologies, indeed the toleration of conflict and disagreement

reflects the very spirit of liberalism. I have also argued that the challenge of liberalism is to resolve such conflicts without suppressing the conditions for individual disagreements: diversity of opinion is, in a sense, what liberal morality is all about.

It is exactly because liberalism must find solutions for (while tolerating) reasonable disagreements that I believe it must necessarily rest upon a commitment to scepticism. If liberalism endorses the toleration of disagreements and conflict, then it must also imply a need for diversity of conceptions of the good: in other words, if there is no diversity, there is no disagreement. I believe that endorsing scepticism is what makes it possible for diversity of views to develop, and for liberalism to flourish. It is only after we accept that our views can be challenged, and that our conceptions of the good require a justification, that we accept the possibility that there may be disagreement and conflict.

The most common way of understanding the challenge to justify our beliefs is that our views are challenged from outside (i.e. my view 'A' is challenged by your view 'B'). I find this conception unnecessarily narrow, in fact before outsiders have any impact on us we must first endorse a frame of mind that makes us open to challenges. It is this frame of mind that encompasses what I have called scepticism, and which is the basis of our moral motivations. Consider the following example: Adolf, a person who displays an illiberal stance, differs from Beth, a liberal person, on the grounds that Adolf will refuse to dialogue with his adversaries, or will do so only out of politeness rather than with the potential

prospect of altering his views. To the extent that Beth believes that her views require justification, hence they are open to challenges, Beth is endorsing scepticism. It is Beth's scepticism of her own position (that is to say, the assumption that her beliefs need to be justified) that determines her liberal nature.

So far I have argued that scepticism and liberalism are closely related. My next step is to argue that the contractualist theory of reasonable agreement is endowed with the sort of scepticism I have endorsed in this chapter, hence it has all the credentials to provide a response to the challenge facing liberalism. I will also argue that other contractualist theories, such as Gauthier's, do not endorse this concept of scepticism, hence it fails to give an adequate answer to the challenge facing liberalism.

Once we accept the definition of scepticism as the doctrine which challenges us to justify our beliefs, then the correlation between scepticism and reasonable agreement should become apparent, since both scepticism and reasonable agreement share a commitment to the justification of our beliefs.

You will recall in fact that following Scanlon I argued that the contractarian theory of reasonable agreement is based on an impartial moral motivation, where one is moved by a desire to justify one's actions or beliefs to others on grounds those others could not reasonably reject. I want to focus for a moment on the idea of 'justification', which I believe is central to the moral motivation championed by Scanlon.

The only way we can explain the desire to justify our actions is if we first swallow a large dose of scepticism, more specifically the sort of scepticism that I have been defending throughout this chapter: as a challenge to justify our beliefs. Of course there is more to Scanlon's idea of reasonable agreement than the endorsement of scepticism, since according to Scanlon we don't want simply to justify our actions as best we can, but we want to do so on grounds that others could not reasonably reject. I will discuss Scanlon's contractualism in more detail in Chapter 3, where the focus will be on the idea of reasonableness and reasonable refutability. The point I want to stress here is that Scanlon's idea of contractualism as reasonable agreement is compatible with scepticism: it is the idea of 'justification' that is common to scepticism and to the contractualist theory of reasonable agreement.

I believe this is an important point since it is this element of scepticism that forms a conceptual bridge between the challenge facing liberalism and reasonable agreement as a response to this challenge. Furthermore the other major neo-contractualist political theory, Gauthier's agreement as rational bargaining, fails to endorse the sort of scepticism that I have been defending, namely, the doctrine which challenges us to justify our beliefs.

Gauthier's agreement echoes Hobbes's theory to the extent that the parties at this stage are not moved by moral inclinations, but merely by a rational desire to pursue their ends as best they can. The agreement itself is characterised by bargaining, where everyone

brings to the bargaining table their natural talents. The process of bargaining of course may at first appear as a materialization of what I have been referring to as the challenge facing liberalism; if conflict is at the heart of ethical liberalism, then Gauthier's contractarianism would seem to be in line with such doctrine, since conflict is also at the heart of Gauthier's idea of agreement. Yet often first appearances are deceiving, and the case of Gauthier is not an exception. The difference between ethical liberalism and Gauthier's contractualism is that while the conflict evoked by ethical liberalism acts upon our moral motivations by forcing us to justify our moral intuitions to others, hence pushing us towards the view that all interests matter intrinsically, the conflict in Gauthier's agreement is devoid of an ethical dimension hence it is nothing more than a show of muscle.

There is no reason why the parties in Gauthier's agreement should endorse the scepticism I have been defending, in fact they could hold any set of beliefs, even the most dogmatic and intolerant. That is to say, the parties in Gauthier's agreement are not challenged to justify their beliefs, no matter how offensive or illiberal these may be. The only type of challenge present in Gauthier's contractualism is between different points of views, where each participant will try to enforce his or her perspective on others. In Gauthier's contract, the parties around the bargaining table are not required to justify their conceptions of the good, instead 'Might is Right' is the natural conclusion of Gauthier's contractualism, even if I suspect Gauthier would never

acknowledge this.

Gauthier's only possible response to the accusation that his theory is reducible to the slogan 'Might is Right' is to invoke a set of individual rights which are not open to bargaining, hence putting limits on someone else's bargaining powers. In the words of Gauthier: "Rights provide the starting point for, and not the outcome of, agreement" (Gauthier 1986, p.222). I don't think this idea of primordial rights can rescue Gauthier, instead as Percy Lehning (1993) points out, it raises the question of whether Gauthier's theory is in fact contractarian. From a contractarian point of view, Gauthier's operation is illegitimate, since the appeal to arbitrary rights is unjustified unless they are also the product of the agreement. Yet this is not the case with Gauthier, since the set of rights he appeals to are formulated prior to the agreement, instead of being the subject of contractarian agreement.

1.VI. Conclusion.

In this chapter I have argued that liberalism is a distinctive theory of political morality. Liberalism has a tradition which links all liberal thinkers over the centuries, indeed what Locke, Mill, Rawls etc. all have in common is the fact that their theories are an attempt to address one central question: how can individual liberties coexist without violating one another. I have referred to this question as the challenge facing liberalism.

There are two things worth pointing out about the challenge facing liberalism. First, that the answer to this challenge must be sought in the idea of egalitarianism, since a precondition for any solution must be that all parties taking part in the liberal enterprise consider the freedom of others to be worthy of equal concern and respect.

Secondly, it is important to point out that although I have called the fundamental liberal question the 'challenge facing liberalism', the term challenge should not be interpreted as a threat, as for example when a boxer challenging the WBC champion is a threat to the latter's title. Instead by challenge here I have in mind a pursuit for something highly valued, the way in which climbing a mountain or writing a doctorate thesis is a challenge. According to this latter understanding of the term, a challenge cannot be divorced from the enterprise of the activity. Thus if I could pay for my Ph.D. degree, or fly an helicopter to the peak of the mountain, I would be undermining the very reason why I choose to undertake these projects.

The reason why it is important to point out the meaning of the term 'challenge' in the context of a discussion of the challenge facing liberalism, is because the conflict between freedoms that liberals ought to be addressing, far from being a threat to liberalism, is its very essence: the political morality of liberalism thrives on conflict. One of the pillars on which liberalism was founded is the belief that disagreements will induce to a better state of affairs.

If we accept these assumptions concerning liberalism as valid, I believe we must also acknowledge that scepticism plays a crucial role in the liberal project. In this chapter I have defined scepticism as the challenge to justify our beliefs, and I have argued that this conception of scepticism is intrinsic to the liberal project. In fact it seems to me that we cannot endorse conflict and disagreement unless we take on board this condition of scepticism. The point is that there is a correlation between the challenge to justify our beliefs, and the challenge facing liberalism: it is only after we accept that our beliefs can be challenged, and therefore require justification, that the conditions for diversity and potential disagreements (the challenge facing liberalism) are set. Thus the challenge facing liberalism, far from being a threat to liberalism, is what liberal morality is all about.

If follows that in order to address the challenge facing liberalism, we must subscribe to a theory which has two features: it must have an adequate conception of egalitarianism, and it must be compatible with scepticism. I believe the first conditions rules out utilitarianism and libertarianism¹⁹, while leaving contractualism as a possible solution. Of course there are many

¹⁹ - While utilitarianism has a conception of equality, as I argue in Appendix 'A' this is inadequate. On the other hand, whether libertarianism endorses a conception of equality is debatable. Kymlicka claims that "the more basic notion of equality [i.e. the idea of treating people 'as equals'] is found in Nozick's libertarianism as much as in Marx's communism" (Kymlicka 1990a, p.4), yet for an alternative view of libertarianism which strongly opposes equality, see Letwin (1983) and Minogue (1989).

different theories of contractualism, hence we need a criterion to distinguish between different types of contractualisms; this is where I think the second condition (scepticism) becomes important. In fact I believe that Scanlon's theory of reasonable agreement encompasses the idea of scepticism as defined above, while Gauthier's contractualism does not.

Of course everything I have said in this chapter was not meant to be a conclusive argument, but was only intended as a general introduction to the arguments I will be defending in the course of this thesis: in the chapters that follow I will address in more detail the concept of reasonable agreement, attempting to show how this can be built into a comprehensive political theory capable of addressing the challenge facing liberalism.

PART I - REASONABLE AGREEMENT IN THEORY

Thomas Nagel maintains that "It is difficult to argue for the intrinsic social value of equality without begging the question" (Nagel 1979, p.106). While this claim has some appeal, it fails to explain why 'equality' remains one of the most debated concepts in contemporary political theory, especially among those who accept the egalitarian ideal. It appears therefore that it is not sufficient to come down on the side of equality; it is also necessary to specify what kind of equality one has in mind. As Kymlicka explains: "If each theory shares the same 'egalitarian plateau' ... [then] the fundamental argument is not whether to accept equality, but how best to interpret it" (Kymlicka 1990a, p.5).

In the previous chapter I argued that acknowledging the equality between social agents is a precondition for solving the challenge facing liberalism, that is, for establishing the conditions for an harmonious coexistence of freedoms. It follows that any liberal political theory must necessarily shed some light on the concept of 'equality'. The contractualist theory of reasonable agreement, which I endorse in this thesis, is not the first theory to respond to the liberal challenge, nor does it

present the only interpretation of equality. For example utilitarianism, by proposing the egalitarian principle 'everyone counts as one, no one more than one', provides a strong response to the liberal challenge.

If reasonable agreement is to legitimately dethrone utilitarianism it will be necessary to show that it promotes a more attractive egalitarian principle than utilitarianism, hence a better response to the liberal challenge. I believe the concept of impartiality is the essence of the egalitarian principle behind reasonable agreement, and that by endorsing impartiality the theory of reasonable agreement responds to the challenge facing liberalism.

The three chapters in Part I will explore the concept of impartiality championed by reasonable agreement. I believe two ideas form the basis of this egalitarian principle: the idea of agreement and the concept of reasonableness. Indeed these two ideas combine to give a different and more attractive response to the liberal challenge compared to utilitarianism.

It is widely acknowledged that Rawls's A Theory of Justice is the single most important fact behind the revival of the social contract tradition. Reasonable agreement emerges from the recent neo-contractualist body of literature, hence it is indebted to Rawls's political theory in general and his contractualism in particular. The merits and limits of Rawls's social contract will be discussed in Chapter 2; more specifically, this chapter will argue that Rawls's contractualism is marred by an internal

contradiction, which threatens to undermine the coherence of his theory of justice. This contradiction is caused by the fact that Rawls endorses conceptual elements from two separate and conflicting social contract traditions: Kant's and Hobbes's social contract theories.

It seems to me that the type of contractualism put forward by Thomas Scanlon (1982) is the best attempt so far to overcome the limits of Rawls's theory while working within the contractarian parameters set-up by Rawls. It is as part of Scanlon's social contract that reasonable agreement makes its debut. The achievements (and shortcomings) of Scanlon's theory will be discussed in Chapter 3.

Notwithstanding the prominence of Scanlon in the literature on reasonable agreement, I believe there are some unresolved problems in Scanlon's theory; failure to solve these problems is the biggest hindrance on reasonable agreement to develop into a comprehensive political theory strong enough to challenge utilitarianism. In particular, I believe Scanlon's theory is not specific enough on the relationship between the two concepts which form the backbone to the theory of reasonable agreement, namely, the *idea* of agreement and the concept of reasonableness. The way in which the concept of reasonableness combines with the *idea* of agreement will be the subject of Chapter 4.

2. RAWLS'S TWO SOCIAL CONTRACTS

Writing on Rawls's political philosophy in general, and his contractualism in particular, twenty years (and several million words of commentary) after the publication of A Theory of Justice, demands an explanation, if not an apology. The reason for embarking on the present analysis of Rawls's contractualism stems from the presentiment that as scholars increase their pile of published material based on Rawls's theory of justice, there is a risk that we gradually loose sight of Rawls's original intention.

I believe that Rawls is partially to blame for this unfortunate state of affairs. The continual reassessment of his theory to accommodate the many critics coming from all sorts of angles, has meant that Rawls has conceded so much ground to his challengers that from where we stand today it is becoming increasingly difficult to recall why Rawls made such an impact on political theory, and what his initial intentions were²⁰. The publication of Political Liberalism and his Amnesty Lecture in 1993 is further evidence that Rawls is now operating on a different

²⁰ - Jeffrie G. Murphy rightly points out that "Even John Rawls, with each new essay, appears more relativistic in his account of morality and of law to the degree that law depends upon or enshrines morality" Murphy (1988), p.240.

wavelength than in 1971, hence in attempting to give the best interpretation of Rawls's political theory we must do without the benefit of Rawls's later insights.

On the question of Rawls's 'intended project'²¹ there are two schools of thought. On one side there are those who argue that Rawls was interested in deducing principles of justice from non-moral premises. I will refer to this first camp as the 'rationalist' camp²². Its supporters focus on Rawls's "original position", in particular the use of rational choice theory and maximin rule, and they tend to see Rawls as a hard-core rationalist. The other camp focuses on those parts of Rawls's work that portray the author's preoccupation with moral intuitions, as embodied in the concepts of "a sense of justice", "reflective equilibrium" or "the well-ordered society", while neglecting Rawls's experiment with rational choice theory. I will refer to this other camp as 'moral-intuitionist'²³.

My impression is that both camps are tapping into real aspects of Rawls's theory, indeed in Rawls's theory we find both rational choice theory and an appeal to our sense of justice, both the original position and reflective equilibrium. It is because of this irreducible tension between Rawls's idea of morality (the sense of

²¹ - Throughout this chapter by Rawls's intended project I will refer to his original intentions and aspirations in writing A Theory of Justice.

²² - The best known exponent of this camp is D. Gauthier (1986).

²³ - This camp includes, among others, R. Dworkin (1978), W. Kymlicka (1990a), M. Sandel (1982).

justice) and the moral decision procedure he set up to justify it (the original position) that there is still much confusion concerning Rawls's intended project.

In this chapter I will argue that the analytical key to Rawls's intended project, and to its intrinsic tension in particular, lies in Rawls's contractualism. In fact although it cannot be disputed that one of Rawls's major contributions to contemporary political philosophy was to revive social contract theory as an alternative to utilitarianism²⁴, it can also be shown that the confusion regarding Rawls's intended project can be traced back to the inadequate nature of his contractualism. The inadequacy lies in the fact that Rawls's contractualism, at the heart of his intended project, is an uncomfortable and unworkable compromise between two incompatible positions in the social contract tradition: the Kantian and Hobbesian social contracts. Indeed Rawls's moral-intuitions (the sense of justice) reveals an affinity with Kant's social contract, while his moral decision procedure (original position) reveals an affinity with a Hobbes's social contract.

In order to see the double nature of Rawls's contractualism, I suggest we start by exploring where Rawls's contractualism fits in relation to classical accounts of the social contract. In 2.I, the general idea of a social contract will be introduced, while in 2.II the different types of social contract argument will be outlined along a continuum line with Kant and Hobbes at the two

²⁴ - See N. Daniels (1975), and C. Kukathas & P. Pettit (1990).

opposing extremes.

Parts 2.III and 2.IV will explore Rawls's account of two key concepts in any social contract theory, namely, social cooperation and agreement. From this account of Rawls's contractualism it will emerge that his approach combines aspects taken from both the Kantian and Hobbesian social contract. Yet the social contract theories of Kant and Hobbes portray mutually exclusive views of morality. The views of Kant and Hobbes on social cooperation and agreement defy reconciliation, with the result that Rawls's contractualism, by attempting such reconciliation, is consumed by an internal contradiction which threatens to jeopardise his intended project.

Echoing Brian Barry's (1989b) claim that we can extrapolate two theories of justice from Rawls's work, a theory of justice as impartiality and a theory of justice as mutual advantage, Part 2.V will show that there are two separate and opposing social contract theories in Rawls's work, and therefore two opposing theories of justice.

In the concluding part of this paper (2.VI) I want to recommend a way in which the tension in Rawls's intended project can be solved, and suggest what I believe is the appropriate interpretation of Rawls's intended project. Indeed some critics have diagnosed the problem with Rawls's theory of justice by reducing all evils to his contractualism, suggesting that Rawls ought to do away with the idea of contractualism tout court, indeed his theory of justice would improve as a result. My suggestion is

to take the opposite line and argue that the problem with Rawls's theory is that Rawls makes only a limited use of contractualism; what Rawls's theory of justice needs is more contractualism, not less.

2.I. The Anatomy of a Social Contract.

Before addressing the more difficult question concerning the anatomy of a social contract theory, we must first agree on a definition of the social contract. Defining the nature and scope of a social contract is more controversial than it may at first appear, in fact we find that different definitions of the social contract are inevitably partial towards certain theories or interpretations of the contractarian enterprise. Furthermore, as we shall see later, in the history of the social contract tradition there are varying and irreconcilable positions.

Consider the following two definitions of the social contract:

- (1) Fundamental to the contractarian enterprise is the task of showing how in appropriate circumstances social order can arise from the operation of individual rationality. (Weale 1993, p.75)
- (2) A social contract theory, in the traditional or 'hypothetical' sense, is one which deduces moral principles for society from

what individuals motivated by self-interest agree to or would agree to. (Lessnoff 1990, p.15)

To the extent that they reflect the views of some of the historical figures behind the contractarian ideology (principally Hobbes, Locke, Rousseau or Kant), the above two definitions are correct. Nevertheless, I feel that these two definitions suffer from being too narrow and partisan.

In the case of Weale, we are told that the social order²⁵ can be reduced to the operation of individual rationality. First of all, it is not clear what Weale means by 'operation'. Since the social contract is not an historical event, but an hypothetical construction, the point is not so much the operation of individual wills as much as the grounds on which individuals could hypothetically agree to cooperate. Secondly, Weale defines 'individual rationality' in terms of individuals pursuing their own interests. Again this may be misleading, in fact if we replace 'individual rationality' with 'individuals pursuing their own interests' we reach the following definition: contractualism is about how the social order can arise from the operation of individuals pursuing their own interests. What is interesting to note is that Weale's definition of the social contract has a striking similarity to an idealized laissez-faire market system. I

²⁵ - "The idea of a social order may be understood as a set of rules and practices within which individuals can pursue their own ends whilst respecting the rights of others to pursue their own ends". Weale (1993), p.75.

should emphasize here that it is not my intention to criticise Weale's understanding of the contractarian enterprise, instead I simply want to draw the attention to the fact that Weale's definition has a marked economic undertone²⁶, which is in contrast with contractualism understood as part of moral theory.

Consider now Lessnoff's definition. He clearly states that the contractarian enterprise entails deducing moral principles from what self-interested individuals would agree to. Lessnoff's definition seems to have strong similarities with Weale's, and I would argue it has the same limitations. In fact Lessnoff's definition excludes the possibility that moral principles are not deduced from the agreement but are an *a priori* assumption of the agreement itself (Kant), or that the social contract is not exclusively concerned with individuals motivated by self-interest, since other forms of motivation may be appealed to²⁷.

In order to avoid the problems which stem from narrow and partisan interpretations of the social contract, I suggest we start from the most general and (I hope) uncontroversial definition of the contractarian enterprise. I suggest the following definition: *A social contract is an agreement based on the consent of every*

²⁶ - In fairness to Weale, it should be said that he discusses the contractarian enterprise in relation to the question of economic justice.

²⁷ - For example, a motivation to act impartially; see Scanlon (1982). Scanlon's theory will be discussed at length in Chapter 3: "Contractualism: Metaphysical not Political".

individual to regulate the benefits of social cooperation²⁸.

The advantage of this definition is that the two key terms, namely agreement and social cooperation, are undetermined, hence the above definition cannot be charged with partisanship.

Now that we have a working definition of the social contract, I suggest we analyze in more detail its anatomy. The formal structure of the contractarian argument follows a two-stage process: in Stage 1 we have the agreement based on individual consent, in Stage 2 we have the benefits of social cooperation. The format of traditional social contract argument can be represented schematically as follows:



Figure 2.1

Remember that this diagram represents only the basic format of the social contract. In fact there are many variations of it, as

²⁸ - In formulating this definition of the contractarian enterprise I have found two recent articles by Samuel Freeman very helpful: "Social contract views work from the intuitive idea of agreement. The appeal of the notion lies in the liberal idea that cooperation ought to be based in the individuals' consent and ought to be for their mutual benefit. Social contract views differ according to how the idea of agreement is specified" (Freeman 1990, p. 122); "Characteristic of the social-contract views are the ideas that social rules and institutions are to be freely acceptable to all persons bound by them, and for their mutual benefit" (Freeman 1991, p.282).

different theories adopt the contractarian framework for different purposes, give different interpretations of its two key moments ('agreement' and 'social cooperation') as well as different accounts of the relationship between these two key moments. Nevertheless all social contract theories, qua contractarian theories, agree that 'agreement' and 'social cooperation' are intrinsic to the social contract.

A growing number of scholars have become convinced that in the history of the social contract tradition, Kant and Hobbes stand on opposite ends of the spectrum. As Will Kymlicka (1991) points out:

There are two basic forms of contemporary social contract theory..... One approach stresses a natural equality of physical power, which makes it mutually advantageous for people to accept conventions that recognize and protect each other's interests and possessions. The other approach stresses a natural equality of moral status, which makes each person's interests a matter of common or impartial concern.... I will call proponents of the mutual advantage theory 'Hobbesian contractarians', and proponents of the impartial theory 'Kantian contractarians' for Hobbes and Kant inspired and foreshadowed these two forms of contract theory. (Kymlicka 1991, p.188)

Analogous accounts to Kymlicka's can also be found in Hampton (1991, p.33), and Lessnoff (1990, p.15).

I will adopt this dichotomy between Kant and Hobbes as the starting point for our enquiry on Rawls's social contract, since it is only by comparison to extreme ideal-types that Rawls's theory can be evaluated. Thus before introducing Rawls, a closer inspection of Kant's and Hobbes's understanding of the social contract is required.

2.II. Kant and Hobbes on the Social Contract.

According to Kant's social contract, the most crucial moment in the contractarian procedure occurs at the stage of the agreement (Stage 1). This is because Kant's concept of 'agreement' is morally loaded, that is to say, it is an agreement based on pre-determined moral motivations. The agreement comes prior to and therefore determines the terms of social cooperation, or in other words the terms of social cooperation are determined by *a priori* moral claims.

Kant's most clear account of the social contract comes from his essay "On the Common Saying: 'This May be True in Theory, but it does not Apply in Practice'", where he argues that if theory is of little practical use, it is not the fault of the theory; the fault is that there is not enough theory. In this essay Kant is concerned to show that in matters of morality any worry about the empty ideality of theory completely disappears, although he also argues that applying the maxim that gives the title to his essay to

matters of morality does very great harm. In the essay Kant deals with the relationship of theory and practice in three separate areas: morality, politics and the cosmopolitan political sphere. The second of these ('On the Relationship of Theory and Practice in Political Right') is a direct attack on Hobbes's De Cive, and it is here that Kant discusses his view of the social contract (and how it differs from Hobbes's).

Kant distinguishes between two types of social contract: as the basis of a society (*pactum sociale*) and as a basis of a civil state, i.e. a commonwealth (*pactum unionis civilis*). The former is more general, and it refers to a union of many individuals for some common end which they all share. The latter is more specific, and it refers to a union as an end in itself which they all ought to share. Needless to say that Kant defends the latter type of social contract as pertinent to the civil or political state, and he sees Hobbes as his major adversary, since Hobbes fails to distinguish between a society and a civil state:

The civil state ... is based on the following *a priori* principles:

1. The *freedom* of every member of society as a *human being*.
2. The *equality* of each with all the others as a *subject*.
3. The *independence* of each member of a commonwealth as a *citizen*.

These principles are not so much laws given by an already established state, as laws by which a state can alone be

established in accordance with pure rational principles of external human right. (Kant 1991, p.74; emphasis in original)

The reason why Kant is keen to discredit Hobbes's social contract is the following: he feels that under Hobbes's contract the head of state has no contractual obligation towards the people, while Kant wants to emphasize that the people too have inalienable rights against the head of state. It is because Kant believes in such inalienable rights that his contract is based on *a priori* principles²⁹.

To recapitulate, Kant's social contract is grounded on a *a priori* moral principles, hence it is the unanimous (moral) agreement that determines the type of social cooperation being pursued:

Figure 2.2 (Kantian Social Contract)

Contrary to Kant, under Hobbes's social contract the most salient moment in the contractarian procedure occurs at the stage

²⁹ - "But reason provides a concept which we express by the words *political right*. And this concept has binding force for human beings who coexist in a state of antagonism produced by their natural freedom ... Thus it is based on *a priori* principles, for experience cannot provide knowledge of what is right, and there is a *theory of political right* to which practice must conform before it can be valid." (Kant 1991, p.86; emphasis in original).

of social cooperation (stage 2). This is because Hobbes's concept of 'agreement' is not morally loaded; according to Hobbes individuals in the state of nature are beyond moral consideration. Morality follows from the agreement, it does not precede it.

Hobbes's rejection of all metaphysical foundations of politics is well documented. In lieu of metaphysics in Leviathan Hobbes prescribes a scientific account of politics, hence he defends a mechanistic account of sensation, and the belief that human motivations can be reduced to appetites and aversions. What is important to note here is that such motivations are beyond any moral appraisal, indeed 'good' and 'evil' do not relate to moral qualities of objects themselves, but are the names of what different individuals desire or hate. This implies that Hobbes undermined all arguments based on an objective basis for morality:

But whatsoever is the object of any mans Appetite or Desire; that is it, which he for his part calleth Good: And the object of his Hate, and Aversion, Evill; and of his Contempt, Vile, and Inconsiderable. For these words of Good, Evill, and Contemptable, are ever used with relation to the person that useth them: There being nothing simply and absolutely so; nor any common Rule of Good and Evill, to be taken from the nature of the objects themselves. (Hobbes 1651, p.24; all emphasis omitted)³⁰

³⁰ - All page references to Leviathan are to the pagination of the original edition of 1651 which is indicated [in square brackets] in the edition by C.B. Macpherson (1968).

Hobbes's denial of moral motives and moral standards is reflected in his accounts of the state of nature and the laws of nature. Concerning the state of nature, Hobbes points out that this hypothetical state is once again beyond moral evaluation:

To this warre of every man against every man, this also is consequent; that nothing can be Unjust. The notions of Right and Wrong, Justice and Injustice have there no place. Where there is no common Power, there is no Law: where no Law, no Injustice. (Hobbes 1651, p.63)

If Hobbes's account of a pre-moral state of nature is surprising, his account of natural laws is even more striking. It should be pointed out that in the 17th century the language of natural laws often implied moral imperatives, especially in the works of those following in the Christian and Aristotelian tradition. Hobbes adopts the notion of natural law, but strips it of all moral connotations. The first two laws of nature are nothing more than necessary conditions if men are to leave the state of nature by the making and keeping of covenants, indeed Hobbes goes as far as saying that it is misleading to refer to these as laws:

These dictates of Reason, men use to call by the name of Laws; but improperly: for they are but Conclusions, or Theorems concerning what conduceth to the conservation and defence of themselves. (Hobbes 1651, p.80)

As Hampsher-Monk rightly points out, for Hobbes the laws of nature are nothing more than rules, and "these rules are advisory, they are not moral absolutes, nor commands" (Hampsher-Monk 1992, p.31).

If the state of nature is beyond moral appraisal, and the laws of nature are not moral imperatives, it follows that (unlike in the case of Kant) there is no right and wrong prior to human agreement. Similarly justice is determined by human agreement, not by moral absolutes prior to such agreement:

For where no Covenant hath preceded, there hath no Right been transferred, and every man has right to every thing; and consequently, no action can be Unjust. But when a Covenant is made, then to break it is Unjust; And the definition of INJUSTICE, is no other than the not Performance of Covenant. And whatsoever is not Unjust, is Just. (Hobbes 1651, p.71; all emphasis omitted)

To recapitulate, according to Hobbes the 'agreement', being stripped of moral connotations, can only be grounded on the bargaining of individual men, and the terms of social cooperation on the idea of mutual advantage. It is the idea of mutual advantage that ensures that the agreement is unanimous:

Agreement as ----> social cooperation ----> sense of bargaining for mutual advantage morality

Figure 2.3 (Hobbesian Social Contract)

In the last analysis, what distinguishes Kant from Hobbes is their different conception concerning the nature of morality: according to Kant morality is intuitively discernible, hence it is invoked prior to the agreement, while for Hobbes morality is evoked after social cooperation. As Hampton rightly points out, according to Hobbes: "morality is a human-made institution, which is justified only to the extent that it effectively furthers human interests" (Hampton 1991, p.36).

So far I have described two ideal-typical interpretations of the social contract idea. I have tried to show that the difference between Kant and Hobbes's contract theories can be seen in terms of three interrelated features: first, their different conception of the key stages of 'agreement' and 'social cooperation'; second, their different understanding of the relation between 'agreement' and 'social cooperation'; and third, their different ideas of the nature of morality. If we place all the social contract theories along a continuum, with Kant and Hobbes at the two extremes, where does Rawls's neo-contractualism lie? In what follows I will argue that in Rawls's contractualism we find elements from both Kant and Hobbes, and that these elements are incompatible with each other. In order to support this claim, I will examine Rawls's conceptions of social cooperation (2.III) and agreement (2.IV).

2.III. Rawls on Social Cooperation.

The idea of social cooperation is indispensable for Rawls's theory of social justice. As he explains at the outset of A Theory of Justice, principles of social justice must necessarily assume social cooperation:

[principles of social justice] provide a way of assigning rights and duties in the basic institutions of society and they define the appropriate distribution of the benefits and burdens of social cooperation. (Rawls 1971, p.4; see also p.7)

In other words, if there is no social cooperation, there is no point drafting principles of social justice. Having established this, what is of interest to us here is Rawls's specific conception of social cooperation.

In what follows I will argue that in A Theory of Justice Rawls appeals to two separate but related conceptions of social cooperation: (a) as promoting mutual advantage and (b) as endorsing a well-ordered society. In Rawls's theory these two conceptions of social cooperation are tightly knit together, to the extent that they become indivisible: "In [justice as fairness] we think of a well-ordered society as a scheme of cooperation for reciprocal advantage" (Rawls 1971, p.33). What Rawls fails to see is that the concepts of mutual advantage and of a well-ordered society reflect

two radically different theories in the social contract tradition, and they also portray two radically different conceptions of morality.

These two conceptions of social cooperation find support in the two schools of thought I referred to earlier, hence the rationalist camp has given prominence to social cooperation as mutual advantage, while the moral-intuitionist camp has focused on the idea of a well-ordered society. I suggest we start by considering first the rationalist camp.

The idea of social cooperation as mutual advantage endorses the principle of mutual benefit; all those who participate in the cooperative venture will benefit from it. It is not surprising to find that the idea of social cooperation as mutual advantage, often upheld by Rawls, is also advocated by neo-Hobbesian political theorists, which explains why Gauthier (1986) is happy to echo Rawls regarding the issue of social cooperation³¹. This similarity between Rawls and Gauthier is misleading, in fact although Rawls adopts the idea of social cooperation for mutual advantage, as we shall see the reason for doing so are very different from Gauthier's.

Samuel Freeman (1990) has made an interesting attempt to isolate Rawls's conception of social cooperation from Gauthier's by focusing on their respective notion of mutual advantage. Freeman points out that while Rawls and Gauthier share the idea of social

³¹ - D. Gauthier (1986), esp. pp.10-13.

cooperation for mutual advantage³², they differ in the characterization of this basic idea: according to Gauthier, cooperation for mutual advantage involves no irreducible moral elements, instead the only valid conception of cooperation is one of efficiently coordinated activity for each person's benefit³³. On the other hand, Rawls's idea of cooperation as mutual advantage endorses irreducible moral notions, which implies that Rawls's conception of social cooperation has a dual aspect:

[I]n addition to a conception of each individual's rational good, the idea of social cooperation has an independent moral component (characterized in Rawls by the notion of fair terms and what is reasonable...). (Freeman 1990, p.124)

While I share Freeman's view that there is a difference between Gauthier and Rawls on the issue of social cooperation, I think Freeman is wrong to look for this within the notion of mutual advantage. That is to say, I don't think there is any difference in the way Gauthier and Rawls understand mutual advantage. Instead, the difference between Gauthier and Rawls is that social cooperation for mutual advantage is Gauthier's only conception of social cooperation, while Rawls appeals also to another conception of social cooperation, a moral conception, enclosed in the idea of

³² - "Both [Rawls and Gauthier] take the idea of reciprocity - the idea that social cooperation should be for mutual advantage - as fundamental" (Freeman 1990, p.123).

³³ - See Freeman (1990) p.124.

a well-ordered society - not surprisingly the concept of a well-ordered society is absent from Gauthier's moral theory. If the idea of mutual advantage is Rawls's first conception of social cooperation, his account of the well-ordered society makes up his other conception of social cooperation.

Another way of stating the difference between Gauthier and Rawls on the question of social cooperation is to say that Gauthier is interested in the *benefits* of social cooperation, while Rawls is interested in the *terms* of social cooperation. In the words of Rawls:

Social cooperation is not merely coordinated social activity efficiently organized for some overall collective end. Rather, it presupposes a notion of fair terms of cooperation. (Rawls 1982; p.164)

Justice as fairness starts from the idea that society is to be conceived as a fair system of cooperation. (Rawls 1985; pp.232-233)

The *terms* of social cooperation are the object of Rawls's second conception of social cooperation, namely, as a well-ordered society. Rawls explains that a well-ordered society is effectively regulated by a public sense of justice, which implies that its members have a strong and normally effective desire to act as the

principles of justice require³⁴. Rawls's idea of a well-ordered society embodies the central theme of his theory of justice, namely, the notion of a sense of justice³⁵. A central aspect of a well-ordered society is the idea of social union, defined as "shared final ends and common activities valued for themselves" (Rawls 1971, p.525), in fact Rawls goes as far as saying that a well-ordered society is itself a form of social union, it is a social union of social unions³⁶.

So far I have argued that Rawls holds two ideas of social cooperation, as mutual advantage and as a well-ordered society. I believe that two conclusions can be drawn from this dual conception. First, that Rawls needs both conceptions for different reasons, that is to say these two conceptions of social cooperation perform different functions in his theory. Thus while social cooperation for mutual advantage is important to Rawls in order to establish the circumstances of justice, the idea of a well-ordered society reflects Rawls's ethical idea of justice. Secondly, these two conceptions of social cooperation reflect two opposing social contract traditions, thus social cooperation as mutual advantage and as a well-ordered society represent respectively Hobbes's and Kant's views of social contract.

³⁴ - Rawls (1971) pp.4-5 and 453-4.

³⁵ - Rawls (1971) explains what he means by a well-ordered society in § 69, which is the first section in Chapter VIII. "The Sense of Justice".

³⁶ - Rawls's concept of social union is emphasized by Sandel (1982) pp.81-2 and 150-1.

It is important for Rawls that these two conceptions of social cooperation are kept separate: in his theory of justice, social cooperation as mutual advantage plays an instrumental role, while the idea of a well-ordered society plays a normative role. The problem with Rawls's dual conception of social cooperation is that he fails to adequately distinguish between their separate functions, therefore we find that the idea of social cooperation as mutual advantage has serious repercussions on Rawls's moral theory.

I believe that the reason why Rawls endorses two separate notions of social cooperation, as well as the reason why he fails to keep these separate, can be traced back to his idea of agreement. It is to Rawls's idea of agreement as original position that I want to turn my attention.

2.IV. Rawls on Agreement.

Rawls's idea of agreement is embodied by the well-known original position, by far the most discussed and criticised aspect of his entire theory of justice. I want to argue that Rawls's notion of agreement is ambivalent, to the extent that it aims to reproduce both aspects of social cooperation (as mutual advantage and as a well-ordered society).

Before pursuing our analysis, it is important to have clear in our minds what the original position represents in Rawls's theory of justice, that is to say, what function the original position

performs in Rawls's theory. The original position is Rawls's idea of a moral decision procedure. As Fishkin points out, moral decision procedures:

define a perspective of impartiality for the equal consideration of relevant claims of interests, and this perspective is offered as the foundation for social choice in a just society. (Fishkin 1984, p.95)

It follows that Rawls's original position can therefore be seen as a device half-way between our general moral intuitions and specific principles of justice. We appeal to the original position in order to clarify and render more precise our general intuitions about equality; in other words the original position is a device which expresses the idea of moral equality (Kymlicka 1990a, pp.68-69) and translates it into principles of justice.

On the basis of this account of the role of the original position in Rawls's theory, we can now analyze the original position in more detail. I believe we can distinguish two features of the original position: (a) its arbitrariness and (b) its internal procedure. I suggest we address these two features separately.

By its 'arbitrariness' I am referring to the fact that the original position is rigged to give certain results. In the words of Rawls:

we want to define the original position so that we get the desired solution. (Rawls 1971, p.141)

The conditions embodied in the description of the original position are ones that we do in fact accept. Or if we do not, then perhaps we can be persuaded to do so by philosophical reflection. (Rawls 1971, p.587)

Rawls admits to modify the original position in order to make sure that it produces the principles of justice Rawls wants to advocate, that is to say principles which match our moral intuitions (Kymlicka 1990, p.67). As Fishkin points out, the original position is only one among many moral decision procedure in contemporary liberal political theory: slight variations in the hypothetical conditions in these choice situations leads to sharply divergent outcomes³⁷.

The second feature concerns the internal procedure or mechanics of the original position. In the original position people choose social principles on the basis of self-interest from behind a 'veil of ignorance'. As Williams correctly points out, this does not mean that Rawls wants to deduce principles of social justice from personal self-interest, instead:

The point is that a self-interested choice in ignorance of

³⁷ – Other examples are Ackerman's notion of "neutral dialogue" and Peter Singer's perfectly sympathetic spectator. See Fishkin (1984), esp. Chapter 4.

one's identity is supposed to model in important respects non-self-interested or moral choice under ordinary conditions of knowledge. (Williams 1985, p.78)

In other words the device of the original position is used to model what would be a *fair* arrangement for people in ordinary life.

Unfortunately for Rawls, there are many problems with the original position as a moral decision procedure. Indeed many critics have rightly argued that the problem with the original position is that rational choice under uncertainty does not represent our moral intuitions on equality: Rawls's idea of social contract, materialised in terms of the original position, fails as a device for embodying a conception of equality³⁸.

I believe we can see exactly why the original position fails to embody our conception of equality by investigating what this idea of rational choice under uncertainty can tell us about social cooperation: what we find is that the assumption of self-interested motivation goes hand-in-hand with the idea of social cooperation as mutual advantage.

It was argued in the previous section that the role of social cooperation as mutual advantage lies in setting the ground for the circumstances of justice, which means that the idea of mutual advantage *per se* is devoid of an ethical dimension. Unfortunately the ethical neutrality of social cooperation as mutual advantage is violated by the original position, in fact Rawls's moral decision

³⁸ - See Williams (1985) p.79 and Barry (1989b).

procedure is characterised by a distinctive ethical dimension, namely, a conception of individual autonomy grounded on rational self-interest.

In order to illustrate this point, it is necessary to recall Gauthier's assessment of Rawls's original position. Gauthier's (1986) main criticism of Rawls's original position is that it excludes bargaining; contrary to Rawls, Gauthier upholds bargaining to the extent that bargaining embodies the spirit of the social contract, i.e. individual consent. If we recall the definition of the contractarian enterprise, the social contract theory refers to the agreement as based on the consent of individuals, thus by interpreting this agreement in terms of bargaining Gauthier is emphasising the individuality of the consent.

The most striking characteristic of Gauthier's conception of agreement is that it is 'agent centred'³⁹, that is to say all reasons centre on the desires and interests of individual agents. According to Gauthier self-interest promotes and reflects the dimension of individuality in individual consent, while bargaining promotes and reflects the idea of consent and agreement.

Gauthier's attempt to force the concepts of individuality and agreement into the narrower concepts of self-interest and bargaining is grotesque: there is more to individuality than the pursuit of self-interest, and bargaining is certainly not the only

³⁹ - I have borrowed this idea from Freeman (1990) p.126.

motive behind an agreement⁴⁰. Having said that, Gauthier's critique can be rescued to point out a potential weakness in Rawls's social contract. If the original position does not involve bargaining or any other type of collective decision process⁴¹, it is not clear under what form individuality appears in Rawls's contractualist framework. Gauthier's conception of agreement as rational bargaining, for all its defects, is strongly agent-centred, therefore it preserves a strong connection to individual practical reason.

It is important to remember that Rawls saw the social contract tradition as an alternative to utilitarianism, which he criticised for not taking account of the separateness of individual persons. This tells us that individuality is important to Rawls, in fact Rawls regards the agreement stemming from the original position as a better alternative to utilitarianism, better in the sense that it respected the individuality of everyone.

I believe it is in order to preserve an element of individuality in his contractualism that Rawls appeals to self-

⁴⁰ - Furthermore Gauthier's idea of agreement as rational bargaining faces the moral problem of justifying social and natural inequalities which greatly determine the outcome of bargaining. Gauthier overcomes this problem by making all natural inequalities legitimate, but of course Rawls has other ideas.

⁴¹ - Rawls tells us that "[T]he original position is not to be thought of as a general assembly which includes at one moment everyone who will live at some time; or, much less, as an assembly of everyone who could live at some time. It is not a gathering of all actual or possible persons". (Rawls 1971, p.139); "[W]e can view the choice in the original position from the standpoint of one person selected at random. If anyone after due reflection prefers a conception of justice to another, then they all do, and a unanimous agreement can be reached" (Rawls 1971, p.139).

interested motivation in the original position. The problem with Rawls's argument is that from a motivation of self-interest he simply cannot deduce the principles of justice he wants. As Brian Barry rightly points out:

But the difficulty that Rawls faces is, I believe, this: there simply is no way in which he can adapt an original position with self-interested choices so as to get it to reflect his basic moral commitments. The foundation is wrong and no amount of work on the detailing can fix the trouble. (Barry 1989b, p.335)

The problem here is that Rawls's appeal to self-interest in the original position has some damaging repercussions for his theory of justice, since the idea of social cooperation for mutual advantage becomes now ethically imbued with a sense of self-interest.

It is true that after 1971 Rawls became aware of the possibility of this interpretation, which he wants to avoid. In fact although in A Theory of Justice Rawls was keen to subordinate his contractarian theory of justice to rational choice theory: "The theory of justice is a part, perhaps the most significant part, of the theory of rational choice." (Rawls 1971, p.16), he later recognised the fallacious nature of this claim:

it was an error in Theory (and a very misleading one) to describe a theory of justice as part of the theory of rational

choice ... What I should have said is that the conception of justice as fairness uses an account of rational choice subject to reasonable conditions to characterize the deliberations of the parties as representatives of free and equal persons; and all of this within a political conception of justice, which is, of course, a moral conception. (Rawls 1985, p.237n)

Clearly Rawls wants to break away from neo-Hobbesian views, and their inclination to integrate morality with rationality and self-interest, yet it seems to me that even his latter restatement concerning rational choice theory is problematic. In fact Rawls still believes that rational choice theory can be used to "characterize the deliberations of the parties as representatives of free and equal persons", while I would argue that the problems in Rawls's theory originate from the employment of rational choice theory *tout court*.

The problem with employing rational choice theory as part of an egalitarian moral discourse is that rational choice theory assumes a motivation of self-interest, and this motivation clashes with the deliberations of the parties as representatives of free and equal persons; at best rational choice theory characterizes the deliberation of free and rational persons, and not of free and equal persons, as Rawls's theory assumes.

It is interesting to notice that in A Theory of Justice Rawls seems to use 'free and equal persons' and 'free and rational persons' interchangeably. Thus Rawls writes that:

[the two principles of justice] are the principles that free and rational persons concerned to further their own interest would accept in an initial position of equality as defining the fundamental terms of their association. (Rawls 1971, p.11; emphasis added).

although the foundation of Rawls's theory is that people should be regarded as free and equal:

[W]henever social institutions satisfy these principles those engaged in them can say to one another that they are cooperating on terms to which they would agree if they were free and equal persons whose relations with respect to one another were fair. (Rawls 1971, p.13; emphasis added)

Judging from these two quotations, Rawls considers free and rational persons to be also free and equal. It would appear therefore that the condition of rationality is instrumental for the equality of a person, thus we are equal to the extent that we are equally rational (or that we have a capacity for equal rationality). Contrary to the view of Rawls, I fail to see the correlation between rationality and equality. The assumption of rationality cannot be a condition of what it means to be treated with equal concern and respect, in fact the claim for egalitarianism is particularly pungent when related to those who do not have a capacity for equal rationality: for example children and

those mentally disabled.

Let's recapitulate the argument so far. I have argued that there are two models for interpreting the device for agreement in Rawls's theory of justice, which are conductive to two opposing conclusions, each relating to opposing ethical theories:

MODEL A: The original position is intentionally rigged in order to yield Rawls's two principles of justice. It follows that these principles of justice reflect a pre-determined idea of moral equality, hence the embodiment of a well-ordered society.

MODEL B: The internal mechanics of the original position give prominence to an individual's motivation for self-interest, which translates into the concept of social cooperation as mutual advantage.

To the extent that Rawls's principles of justice and idea of social cooperation are traceable to a pre-determined moral sense of justice, under MODEL A Rawls is advocating a Kantian social contract theory. At the same time the original position assumes a form of individualism compatible to mutual advantage, which may itself determine our sense of morality, hence under MODEL B Rawls may be seen as advocating a Hobbesian social contract theory.

In Part 2.V below I will enquire what conclusions can be deduced from the claim that Rawls simultaneously upholds two separate and conflicting views of morality and of the social contract.

2.V. Social Contracts and Theories of Justice.

So far I have tried to show that although Rawls advocates a return to the social contract tradition, he is ambiguous as to the kind of social contract theory he wants to defend, in fact both his conception of social cooperation and agreement reflect evidence of two opposing theories of the social contract. On one side Rawls justifies the claim for a well-ordered society (a society effectively regulated by a public conception of justice) on the grounds of a moral sense of justice⁴². Yet on the other hand Rawls upholds social cooperation for mutual advantage based on a sense of rational self-interest. The former feature echoes a Kantian social contract, the latter a Hobbesian social contract.

There is no doubt that Rawls's intention was to formulate a non-Hobbesian social contract theory, as he explicitly stated:

As the text suggests, I shall regard Locke ... Rousseau ... and Kant ... as definitive of the contract tradition. For all of its greatness, Hobbes's Leviathan raises special problems.
(Rawls 1971, p.11n)

So why does Rawls feel that he must endorse the (Hobbesian) idea of social cooperation for mutual advantage? As I explained

⁴² - Indeed Rawls tells us that in a well-ordered society citizens are moved by a sense of justice.

already, Rawls adopts the idea of social cooperation for mutual advantage based on individual self-interest, for two reasons. First of all, because it provides what Hume referred to as the 'circumstances of justice', that is to say, the conditions which establish the usefulness of principles of justice. Secondly, it is in order to retain an element of individuality in his theory of justice; having rejected the idea of compromise and bargaining, Rawls needs to counterbalance his decision to exclude bargaining from his theory by emphasizing the element of individuality in the notion of social cooperation, since individuality is a *sine qua non* postulate for any social contract theory.

Rawls's failure to combine the idea of a well-ordered society, where individual agents are moved by a sense of justice, with that of social cooperation for mutual advantage, where the motivation is one of rationality aimed at advancing one's own interest, has some serious repercussions on Rawls's moral and political theory. Indeed Rawls's contractualism, embodying two conflicting theories of the social contract, is the cause of an unworkable contradiction in his theory of justice. Here I am referring to Brian Barry's well known claim that Rawls endorses two conflicting theories of justice, one pointing to justice as impartiality, the other to justice as mutual benefit.

In the first volume of his tetralogy on social justice, Barry (1989b) argues that we can extrapolate two theories of justice from Rawls's work. The two moral capacities depicted by Rawls, one for a conception of the good (expressed by a rational plan of life),

the other for a sense of justice (expressed by a regulative desire to act upon certain principles of right) are the founding stones of two predominant and conflicting approaches to social justice, respectively justice as mutual advantage and justice as impartiality. According to Barry, these two approaches to social justice are both implicit in Rawls's A Theory of Justice.

It seems to me that the two moral capacities above mentioned, namely the rational pursuit of one's conception of the good and the sense of justice, which Barry traces back to Rawls's theory of justice, are also the same moral capacities assumed by Hobbes and Kant in their respective social contract theories. It follows that the tension within Rawls's A Theory of Justice between two opposing theories of justice is the symptom of a deeper contradiction, namely, between two conflicting social contract traditions: one Hobbesian, the other Kantian.

To recapitulate, I believe that the ambiguity in Rawls's theory of justice stems from his endorsement of two opposing social contract traditions. In other words, in Rawls's theory we find elements of two separate and opposing social contract theories, and therefore elements of two opposing theories of justice. Justice as impartiality stems from a Kantian interpretation of the social contract, while justice as mutual benefit stems from a Hobbesian interpretation.

2.VI. Conclusion.

The aim of this chapter was to shed some light on Rawls's intended project by exploring that aspect of his work that has been acclaimed as Rawls's greatest contribution to contemporary political philosophy: having revived the social contract theory as a response to utilitarian moral and political theory. I believe that establishing the exact role of contractualism in Rawls's A Theory of Justice is the key to a full understanding of the strengths and limits of Rawls's intended project. In other words it is in light of Rawls's commitment to the social contract that his intended project must be understood.

In this chapter I argued that there are elements of two opposing social contract traditions in Rawls's theory: Kant's idea of a morally loaded agreement, and Hobbes's idea of social cooperation for mutual advantage. I also argued that the contradiction stemming from these two traditions is not resolved by Rawls but merely reflected in his theory of justice, in fact it is possible to detect two opposing theories of justice in Rawls's work.

Brian Barry's interpretation of Rawls's work goes a long way towards justifying the argument presented in this chapter that there are elements of two social contract theories in Rawls's work: one Hobbesian, the other non-Hobbesian. The Hobbesian element in Rawls's theory of justice can be found in the idea of social cooperation for mutual advantage and the motivation for self-interest in the original position, while the Kantian element lies in the idea of an intuitive sense of justice and the concept of a

well-ordered society.

Considering the unresolved problems which derive from Rawls's use of the social contract, it is not surprisingly that Rawls's revival of the social contract has received mixed responses. This can be seen in the way in which the 'rationalist' and 'moral-intuitionist' camps have reacted to Rawls's attempt to revive interest in the social contract tradition.

There is little doubt that the rationalist-camp is keen on Rawls's contractualism. This is not surprising considering that the original position is the centre piece of the rationalists' interpretation of Rawls's intended project. By identifying with Rawls's original position and the use of rational choice theory, some exponents of the rationalist camp have seized upon the opportunity to uphold the original position and the logic of rational choice in order to deduce principles of justice. This is the case of David Gauthier, who argues that only a theory of justice as bargaining can result from rational postulates.

On the other hand the moral-intuitionist camp, remaining faithful to Rawls's ethical assumptions concerning fairness, has argued that if the social contract amounts to the original position (i.e. rational choice under uncertainty), then we can do without the social contract tout-court. Moral-intuitionists claim that hypothetical contracts do not supply an independent argument for the fairness of certain terms of agreement, instead it is a fundamental right to equal concern and respect⁴³ that gives Rawls's

⁴³ - What Dworkin calls the 'deep' theory.

theory its justificatory power.

The irreconcilable tension between Rawls's moral intuition and his moral decision procedure has resulted in a on-going debate concerning the correct interpretation of Rawls's political theory. Having analyzed Rawls's contractualism in depth, the time has come to draw some conclusions: What is the role of Rawls's contractualism? Is Rawls's political theory 'rationalist' or 'moral-intuitionist'?

While I agree with the rationalist camp that Rawls's original position does point to a rationalist interpretation, I also think that reducing Rawls's theory of justice to an attempt to deduce moral claims from rational axioms is to misunderstand Rawls's intended project. Indeed if that was the case, then one would need to explain Rawls's claim that the original position is rigged to give certain predetermined results.

Contrary to the rationalist camp, I believe Rawls's intended project was to ground principles of justice on specific egalitarian moral-intuitions, namely, the idea of persons as free and equal, indeed the principles of justice are the best endorsement of this moral assumption. Yet I also believe that Rawls's political theory cannot be divorced from the contractarian approach, since it is by endorsing contractualism that Rawls is able to challenge utilitarianism in the first place. Contractualism made it possible for Rawls to justify his principles of justice by retaining an element of individual autonomy, although in his attempt to do so Rawls unexpectedly and carelessly bridged the gap between his

political theory and Hobbesian contractualism.

In the following chapter, I will argue not only that it is possible to endorse Rawls's egalitarian moral intuitions without abandoning contractualism, but that contractualism is the best method of explaining and justifying egalitarian principles. It follows that the major weakness of Rawls's theory is not his contractualism, but the fact that he makes limited use of contractualism in his theory, or in other words, the problem is that there is not enough contractualism in Rawls's theory.

Rawls appeals to the contract device exclusively to indirectly tease out the implications of certain moral premises or motivations; what he should have done instead is to appeal to contractualism in order to directly justify our moral intuitions.

I am referring here to the well known distinction between ethical and meta-ethical theory: briefly, ethical theory is concerned to determine which actions are right or wrong, while meta-ethical theory is an enquiry into the nature of morality itself (that is to say, in what sense moral judgements can be held to be correct)⁴⁴. Rawls evokes the social contract at the level of the ethical theory, but not at the meta-ethical level. This is problematic since unless contractualism works on both levels there is a risk of embracing contradictory positions between the ethical and meta-ethical levels. Rawls's theory of justice is an example of this contradiction. In order to resolve the contradiction within Rawls's theory, the contractarian device must be raised from the

⁴⁴ - For an analysis of these two notions, see Scanlon (1992).

ethical level to the meta-ethical level⁴⁵. It follows that the problem with Rawls is not the social contract *per se*, but the use Rawls makes of it.

I believe Thomas Scanlon was one of the first to understand this. In a famous article from 1982, "Contractualism and Utilitarianism", Scanlon attempts to explain egalitarian moral intuitions (i.e. our moral motivations) with the help of a contractualist framework. I believe that Rawls's intended project can be rescued if we replace the idea of the original position with Scanlon's idea of a reasonable agreement. The advantage of Scanlon's conception of agreement over Rawls's is that in the former case our moral intuitions are intrinsically related to the contractualist approach. The basic idea of the social contract is the basis of our own moral intuitions: the contractarian model, and our moral intuitions, cannot be divorced. It follows that the social contract device is more than a heuristic apparatus for "teasing out" the implications of our moral intuitions. The concept of agreement reveals the nature of morality, that is to say the social contract is the basis of our sense of justice, and thus consequently of a theory of justice.

⁴⁵ - As will shall see in the following chapter, this is the contribution of Thomas Scanlon to contemporary moral and political theory.

3. CONTRACTUALISM: METAPHYSICAL NOT POLITICAL

In the previous chapter I argued that there are elements of two opposing social contract traditions in Rawls's theory: Kant's idea of a morally loaded agreement, and Hobbes's idea of social cooperation for mutual advantage. I also argued that the contradiction stemming from these two traditions is not resolved by Rawls but merely reflected in his theory of justice, in fact it is possible to detect two opposing theories of justice in Rawls's work.

Some commentators have argued that all the complications in Rawls's theory of justice originate from his contractualism, in fact the original position is at best useless, and at worse counter-intuitive. This general feeling of dissatisfaction is captured by Kymlicka when he writes that:

So the contract device adds little to Rawls's theory. The intuitive argument is the primary argument, whatever Rawls says to the contrary, and the contract argument (at best) just helps express it. But it is not clear that Rawls needs an independent contract argument. (Kymlicka 1990a, p.69)

In this chapter I will argue not only that it is possible to endorse Rawls's egalitarian moral intuitions without abandoning contractualism, but that contractualism is the best method of explaining and justifying the moral intuition of impartiality. In other words, contractualism is particularly important when applied to the meta-ethical level. It follows that the major weakness of Rawls's theory is not his contractualism, but the fact that Rawls does not make use of contractualism in order to explain and justify his intuitive argument.

A theory of justice based on the intuition of impartiality, such as Rawls's, needs the contractualist device in order to explain and justify its meta-ethical root. The meta-ethical level deals with questions relating to the nature of morality, for example in what sense moral judgements can be held to be correct. Contractualism helps us to answer these type of questions, rather than with establishing which actions or practices are right or wrong.

I believe Thomas Scanlon was one of the first to understand that contractualism had a meta-ethical dimension. In a famous article from 1982, "Contractualism and Utilitarianism", Scanlon attempts to explain egalitarian moral intuitions with the help of a contractualist framework, although as we shall see Scanlon's contractualism differs from Rawls's in both its scope and aim. This chapter will therefore focus on the contribution of Scanlon to the recent literature on contractualism.

This chapter is divided in six parts. Parts 3.I and 3.II

comprise of a refutation of the argument, championed by Kymlicka, that it is a major weakness of some contemporary theories of justice to ground the idea of justice as impartiality on the notion of agreement. In order to vindicate the role of contractualism in contemporary theories of justice, I will argue that the idea of agreement is crucial for a meta-ethical discourse. Thus Scanlon's contractualism can be distinguished from Rawls's contractualism to the extent that the former operates at the meta-ethical level, while the latter operates at the ethical level.

Parts 3.III and 3.IV analyzes Scanlon's contractualism in general, and his theory of reasonable agreement in particular. Here I will argue that Scanlon's theory has two major merits, namely challenging utilitarianism at its very moral foundations, and formulating an original contractualist theory which differs from Rawls's on some key issues.

Part 3.V shows that although Scanlon's contractualism represents an improvement on Rawls's theory, there are still some major unresolved problems with Scanlon's theory, for example Scanlon's failure to conceptualise with adequate precision the key idea of reasonableness and his over-reliance on the psychological faculty of 'desiring'.

In order to strengthen the case for a theory of reasonable agreement, I believe these complications must be overcome. Indeed we find that failure to deal with these problems is a major weakness in the theories of Samuel Freeman and Thomas Nagel, two strong sympathizers of Scanlon's theory. The theories of Freeman

and Nagel will be analyzed in Part 3.VI.

3.I. The Idea of Agreement.

If there is a recurring theme in Kymlicka's writings, it must be his severe scepticism of the contractarian device in moral theory. In its milder form, such scepticism is portrayed by comments referring to the redundant nature of contractualism in a Kantian moral theory:

What is not clear is whether the contract device does any work defending or developing these ideas [i.e. basic elements of our everyday moral understanding v.b.]. (Kymlicka 1991, p.193)

In its more austere form, Kymlicka's scepticism is a direct attack against those neo-Rawlsians⁴⁶ whose project is to vindicate a theory of justice as impartiality by appealing to the notion of agreement. On the receiving end of Kymlicka's attack we find Scanlon and Barry.

According to Kymlicka, the notion of agreement is not a valid foundation for a moral discourse. The reason for this is that the parties who can potentially partake in an agreement reflect a

⁴⁶ - By neo-Rawlsians I am referring to those who are intuitively sympathetic to Rawls's A Theory of Justice, but who are frustrated by his line of work since 1980.

narrow spectrum of moral beings, namely "competent adults" (Kymlicka 1990b, p.110). In other words, the notion of agreement unfairly leaves out a range of people who have a right to be considered by a moral theory, for example infants, future generations and the demented⁴⁷:

But what does it mean to desire impartial agreement with infants, or to desire to be able to justify one's actions to people who don't yet exist? If someone is incapable of being a party to an agreement with us, does that mean we lack any moral motive for attending to their interests? The emphasis on agreement within impartiality seems to create some of the same problems that the emphasis on bargaining power creates within mutual advantage theories: some people will fall beyond the pale of morality, including those who are most in need of moral protection. (Kymlicka 1990b, p.110)

Thus while the notion of agreement applies for competent adults, it does not apply beyond this small circle of beings.

It ought to be pointed out that for all its critical sting, Kymlicka's argument is not rejecting the idea of justice as impartiality. The controversy between Kymlicka on one side, and

⁴⁷ - I have left out one entry from Kymlicka's list: 'animals'. It seems to me that the question concerning animals raises special problems that are not encountered in the case of babies, the unborn and the demented. Since it is an issue of contention whether human beings have the same moral duty towards animals as they have towards each other, I feel it is best for the moment to neglect the question concerning non-human beings.

Scanlon and Barry on the other, is not whether it is desirable to defend a theory of justice as impartiality, since they all agree on the necessity to ground justice on the idea of impartiality. Instead the dispute concerns the nature of impartiality: according to Kymlicka we should not ground impartiality on some kind of agreement, as Scanlon and Barry seem to think, since impartiality refers to the belief that the well-being of all moral beings matters intrinsically. In other words according to Kymlicka impartiality simply means giving all interests equal weight:

Some beings with moral status can be given a justification, others can't. What makes them all moral beings is the fact that they have a good, and their well-being matters intrinsically. (Kymlicka 1990b, p.111)

While the above assertion by Kymlicka can be seen as an alternative account of impartiality to Scanlon's, it must be said that Kymlicka's polemic in relation to Barry's work takes an interesting twist, since Barry entertains the idea (so dear to Kymlicka) that impartiality is a criterion that gives all interests equal weight⁴⁸. For this reason I feel that Kymlicka's critical assessment of Barry's theory deserves closer attention.

Kymlicka is all in favour of Barry's theory when the latter claims that the basis of our impartial moral motivation is the recognition of others as having legitimate claims to have their

⁴⁸ - See Barry (1989b), p.269.

interests taken into account. Yet such harmony is short lived; the cleavage between Kymlicka and Barry occurs when Barry tries to connect the demand for agreement with the consideration of people's interests. In the words of Kymlicka:

Unfortunately, Barry invariably goes on to muddy the waters by rephrasing these interest-based statements of impartiality in terms of agreement. But if the aim is that people's legitimate interests be taken into account, then why not just say that impartial theorists try to find principles that give equal weight to everyone's interests? The further claim that everyone has to agree to the principles does nothing except put those who are incapable of agreement beyond the pale of justice. (Kymlicka 1990b, pp.111-112)

In what follows, I want to refute Kymlicka's claim that the notion of agreement 'does nothing' except make things worse. Contrary to Kymlicka's claim, I believe that agreement does something fundamental for a theory of justice, something that Kymlicka totally neglects, namely, it enables us to distinguish between people's interests and people's legitimate interests. It is curious that Kymlicka should not pick up on this point considering that his own criterion of impartiality assumes legitimate interests.

It seems to me that there are two problems with Kymlicka's account of contractualism. First of all, there is a sense in which

an agreement can only be between competent adults living at the same time. After all, one of the basic functions of contractualism is to foster peace and harmony, a view endorsed by Hobbes in the Leviathan and by Rawls (1993b) in his International Amnesty Lecture. Considering this role played by a social contract theory, it is not surprising that the parties taking part in the agreement should be competent adults.

In fact as Scanlon (1988b) argues, the very essence of an agreement is based on the nonmoral capacity for critically reflective, rational self-governance:

Basic to morality as I understand it is an idea of agreement between individuals *qua* critics and regulators of their actions and deliberative processes. Critically reflective, rational self-governance is a capacity which is required in order for that idea not to be an idle one. It follows that moral criticism is restricted to individuals who have this capacity and to actions which fall within its scope. (Scanlon 1988b, p.175)

Competent adults are, by definition, imbued with the capacity for critically reflective, rational self-governance. Hence contractualism cannot be other than between competent adults. It follows that while there is no limit to who or what can be a recipient of morality, from the environment to foetuses, there is in fact only one originator of morality, namely, the competent

adult human being⁴⁹. In other words if a tree or a foetus is promoted to the realm of moral beings, thus enjoying all the benefits this entails (such as protection), it is only because competent adults, being the originators of morality, have agreed that trees or foetuses should be considered moral beings while the flea should not.

Secondly, it seems to me that Kymlicka has a mechanical conception of agreement. What Kymlicka fails to appreciate is that what is important for a theory of justice is not simply the agreement as much as the *idea of agreement*. The crucial element in the notion of agreement for a contractualist approach is the fact that the parties seeking an accord are willing to compromise their conception of the good in order to make room for others. It is this willingness to compromise that is the key to the *idea of agreement*.

It follows that contrary to Kymlicka's view, the idea that everyone has to agree to the principles performs a fundamental role in a theory of justice, namely, it enables us to distinguish acceptable from unacceptable interests. In other words, what makes an interest acceptable or legitimate is the fact that one is willing to compromise (or revise) one's interest in order to seek agreement. It is interesting to note that all the time Kymlicka is assuming that what matters are legitimate interests, not just any interests: "At the deepest level, justice is about the equal

⁴⁹ - I am grateful to Brian Barry for pointing this out to me. I have also found Jeremy Waldron's review article of Dworkin's Life's Dominion particularly relevant on this point. See Waldron (1994).

consideration of our legitimate interests ..." (Kymlicka 1990b, p.112; emphasis added)⁵⁰. Yet Kymlicka fails to tell us when an interest is legitimate and when it is illegitimate. I believe the answer lies in the *idea of agreement*: it is the process of seeking an agreement that legitimizes one's position.

It should also be emphasized that the *idea of agreement*, and the criterion that gives all interests equal weight, are not mutually exclusive, which is why Barry appeals to both in his account of impartiality. We start from the criterion that gives all interests equal weight, but then in order to overcome fundamental conflicts it is necessary to distinguish between legitimate and illegitimate interests. This is where the *idea of agreement* applies.

My critique of Kymlicka's refutation of agreement is echoed in a recent article by Kai Nielsen (1992). Nielsen argues that although there is plainly something right about Kymlicka's argument, he may have been too quick to dismiss the notion of agreement. While Nielsen accepts Kymlicka's line that intuitively justice is about the equal consideration of our legitimate interests, he goes on to ask how do we know that justice is what we think it is. According to Nielsen, the answer to this question lies in the notion of agreement:

How do we know that this is what justice is and that this is

⁵⁰ - "But if the aim is that people's legitimate interests be taken into account" (Kymlicka 1990b, p.112; emphasis added).

what justice requires such that we must act in this way if we would be just and that for there to be just social institutions our social practices must be so structured? It is here that agreement may come in by the back door. (Nielsen 1992, pp.93-94)

According to Nielsen, the notion of agreement is important in order to establish the foundations of our moral claims. The alternative to the notion of agreement seems to be the discredited intuitionist and natural law traditions.

I believe Nielsen is making a crucial point here, which I will investigate further in Part 3.II, namely, that the notion of agreement (and contractualism in general) applies to the fundamental or justificatory level of meta-ethics. According to Nielsen, at a fundamental level the notion of agreement is needed to justify a specific wide reflective equilibrium, just as a wide reflective equilibrium justifies an account of justice as impartiality:

It is not that the substantive principles and claims of social justice are not what Kymlicka says they are or that justice is what we can agree on in certain idealized situations but that, if we are to show that Kymlicka's or anyone else's substantive claims of justice are justified, we must show that there is such agreement. (Nielsen 1992, p.97)

In order to sustain the claim that contractualism applies at a fundamental level, it is necessary to distinguish between the ethical and the meta-ethical level. It is this question that I want to address now.

3.II. Contractualism and Moral Theory.

Arguments in moral philosophy can be distinguished between two levels of inquiry: the ethical and meta-ethical level. At the ethical level we are concerned with judging which actions or practices are right or wrong, while at the 'philosophical' or 'meta-ethical' level we are concerned with the nature of morality, or in what sense moral judgements can be held to be correct⁵¹. The question I want to address now is the role of contractualism in moral theory, in other words, at what level of moral enquiry is contractualism engaged by Rawls and Scanlon.

Rawls's contractual conception is represented by the original position. Yet in the previous chapter I argued that the original position is arbitrarily devised in order to give pre-determined principles of justice. This is an important point since it indicates that the original position does not presuppose some particular view of moral motivation, instead Rawls discusses the relationship between moral theory (of which justice is a part) and

⁵¹ - Scanlon (1992) refers to the former as Moral Enquiry, and the latter as Philosophical Enquiry. In this chapter, I will call the former 'ethical level', and the latter 'meta-ethical level'.

the ordinary moral judgements in his account of reflective equilibrium⁵².

What does this tell us about the role of contractualism in Rawls's moral theory? Basically, that the original position is silent about questions of moral motivation or the nature of morality (questions dealt with at the meta-ethical level). Instead the original position operates exclusively at the ethical level, determining the rightness of principles of justice.

Yet as I have argued in the previous chapter, there is an unresolved contradiction in Rawls's theory between the motivation operating in the original position, and the moral motivation on which the two principles of justice as fairness are based. What Rawls's theory requires is an investigation of the moral motivations assumed by his principles of justice as fairness. Rawls fails to see that contractualism can be used at the meta-ethical level to explain the nature of our moral intuitions. In other words, the foundation of the moral intuition which Rawls is appealing to can be traced back to the idea of agreement.

Appreciating the meta-ethical dimension of contractualism is arguably Scanlon's greatest contribution to contemporary moral and political philosophy. Thus while Rawls's contractualism is present only at the ethical level, Scanlon pushes contractualism beyond the ethical to the meta-ethical level. Scanlon argues that contractualism concerns an agreement between persons moved by a desire to justify their actions to others on grounds they could not

⁵² - See R. Plant (1991), pp. 101 & 107.

reasonably reject. As we shall see later, according to Scanlon this idea of contractualism is the essence of morality.

The two key concepts in Scanlon's contractualism are 'desire' and 'reasonableness'. The desire to find an agreement based on the concept of impartiality is the moral motivation in Scanlon's contractualism. Equally important to this desire is the notion of reasonableness, in fact one does not exaggerate by saying that the condition of reasonableness does all the work in Scanlon's contractualism.

Scanlon addresses the question of defining the notion of the 'reasonable' in an important footnote: "Reasonably, that is, given the desire to find principles which others similarly motivated could not reasonably reject" (Scanlon 1982, p.116n). In the attempt to define the notion of reasonableness, Scanlon introduces the idea of others being "similarly motivated". I believe that the notion of all the parties seeking agreement being "similarly motivated" holds the key to Scanlon's concept of reasonableness.

In fact this simple idea of others being "similarly motivated" hides an important belief, worth exploring in some detail. Why would Scanlon, for example, want to justify his actions to others who are similarly motivated? That is, to others who also want to justify their actions to Scanlon on grounds he could not reject? It is important to be clear on what is meant by "similarly motivated".

To be "similarly motivated" does not mean that we have the same goals, or conceptions of the good. The 'others' in Scanlon's assertion are not people who think like Scanlon, or share his set

of beliefs. In other words it is not a community made up of many Scanlons. Instead in the process of seeking an agreement, the idea of being "similarly motivated" refers to the starting point rather than finishing point. In other words we share certain assumptions concerning, roughly, what conceptions of the good are acceptable and what are not acceptable, while we do not all share the same conception of the good. To this extent the idea of being "similarly motivated" is what makes the agreement unanimous.

It follows that the idea of being "similarly motivated" is related to the other key concept in Scanlon's theory of moral motivation, namely, the desire to justify one's actions to others, or the desire to find principles others could accept. This desire can more appropriately be called a desire for reasonable agreement.

The desire for reasonable agreement is the cornerstone of Scanlon's contractualist account of moral motivation: Scanlon's contractarians desire to cooperate, and the terms of cooperation arise from the agreement itself⁵³. I will have more to say about Scanlon's contractualism shortly, in particular how it differs from other forms of contractualism and why it is superior to utilitarianism. What is important to emphasize for the moment is the fact that Scanlon's contractualism operates at the meta-ethical level.

⁵³ - In the words of John Charvet, Scanlon's agreement is grounded "in the will of persons to pursue their good together as members of a cooperative association on terms which all can accept from an impartial point of view", from John Charvet's "Contractarianism in International Political Theory", unpublished manuscript.

According to Scanlon's idea of contractualism, the parties involved must be moved by a desire to seek an agreement that everyone would find acceptable. Therefore what is important for contractualism is not the agreement itself as much as the motivation behind the agreement. The disposition to find an agreement that everyone would find acceptable implies being open to compromise or revise one's own conception of the good in order to make room for others, assuming that others would do the same for you or anyone else. It is the motivation behind the idea of agreement that explains the nature of morality, or in other words which explains in what sense moral judgements can be held to be correct. The disposition to seek an acceptable agreement is the foundation of liberal and egalitarian meta-ethics, hence contractualism is first and foremost a meta-ethical notion.

In order to see how Scanlon's idea of contractualism operates at the meta-ethical level, it is necessary to give a detailed account of "Contractualism and Utilitarianism", especially as the theory of reasonable agreement I will be presenting in the next chapter uses Scanlon's approach as its foundation.

Scanlon's self-proclaimed goal in this influential essay was to adopt the ideas stemming from the social contract tradition in order to extract a non-utilitarian account of moral motivation. This article has been appreciated mainly for its critique of utilitarianism and for advocating the social contract tradition as an alternative to utilitarianism. Yet I believe there is more to Scanlon's contractualism than has so far been recognized.

In what follows, I will argue that what many have failed to notice (partly because of Scanlon's self-proclaimed goal and partly because his writing is undeniably dense) is that Scanlon's essay is composed of two parts. The first, well documented part is a critique of utilitarianism, while the second part, generally neglected, is a critique of other neo-contractualist theories, notably Rawls's. In other words Scanlon is not simply mounting a critique of utilitarianism, rather he is also advancing an original interpretation of the social contract, which differs from both its classic formulation (Kant and Hobbes), and from contemporary versions of neo-contractualism (Rawls and Gauthier).

I believe that a proper analysis of Scanlon's contractualism requires that these two parts are dealt with separately, even though analytically they are interrelated. In fact Scanlon adopts a contractualist stand as the focus of his critique of utilitarianism, while at the same time using his critique of utilitarianism to distinguish his theory from that of Rawls.

3.III. Scanlon's Critique of Utilitarianism.

Working on the impetus of Rawls's theory of justice, Scanlon succeeded where Rawls had failed⁵⁴: he developed a plausible

⁵⁴ - A more accurate formulation of this claim would be that "Scanlon partially succeeded where Rawls partially failed". Scanlon's success is only partial since he only provided us with a rough outline of a neo-Rawlsian version of contractualism. In the concluding paragraph of his article Scanlon writes "I have

alternative account to utilitarian moral motivation. According to Scanlon, an adequate moral philosophy must make clear to us the nature of morality, or in other words provides us with reasons to act. Providing us with a seemingly persuasive theory of moral motivation is utilitarianism's comparative advantage with respect to many other theories. Hence a non-utilitarian moral philosophy must first show that utilitarian moral motivations are inadequate, and consequently formulate an alternative theory of moral motivation to utilitarianism.

According to Scanlon the moral motivation behind utilitarianism consists in maximising the sum of individual well-being⁵⁵, or in other words it aims to generate maximum aggregate happiness⁵⁶. Scanlon makes a distinction between two types of utilitarianism: philosophical utilitarianism and normative utilitarianism. Although Scanlon finds philosophical utilitarianism attractive (the idea of grounding the nature of morality on individual well-being), he rejects normative utilitarianism, in

described this version of contractualism only in outline. Much more needs to be said to clarify its central notions and to work out its normative implications" (Scanlon 1982, p.128). Although Scanlon's 1988 Tanner lectures on "The Significance of Choice" are an attempt to carry on the work he started in 1982, he has not yet attempted the hard work of exploring the implications of his contractarian moral motivation on a theory of justice. On the other hand Rawls's failure is only partial since justice as impartiality was always central to his work; as Barry (1989b) argues, the elements of justice as impartiality in Rawls's construction are at least as important as the elements pertaining to justice as mutual advantage.

⁵⁵ - Scanlon (1982) p.110.

⁵⁶ - Scanlon (1988a) p.137.

particular he rejects the idea that maximising aggregate well-being is forced on us:

once philosophical utilitarianism is accepted, some form of normative utilitarianism seems to be forced on us as the correct first-order moral theory. (Scanlon 1982, p.109)

What Scanlon finds intuitively wrong with utilitarianism is the coercive and inegalitarian aspect of this theory. Maximising aggregate happiness becomes the independent criterion imposed on each one of us as the basis for reaching agreement *regardless of whether we find the agreement acceptable*.

Scanlon rejects the idea of an enforced agreement, arguing instead for an alternative account of moral motivation along impartial lines. This motivation is triggered by the belief that an action is wrong, in Scanlon's own words:

if its performance under the circumstances would be disallowed by any system of rules for the general regulation of behaviour which no one could reasonably reject as a basis for *informed, unforced general agreement*. (Scanlon 1982, p.110; emphasis added)

Echoing Rawls, Scanlon finds in the social contract tradition the antidote against utilitarianism. Yet this rejection of utilitarianism is only half of Scanlon's project. As I pointed out

earlier, Scanlon is fighting simultaneously two battles. On one front he is criticising utilitarianism, yet on the other front Scanlon is criticising other social contract theories.

Although Scanlon was greatly influenced by Rawls's work, in rejecting utilitarianism Scanlon was not simply making a general statement in favour of the revival of contractualism along the lines suggested by Rawls in 1971. We have seen that Scanlon rejects utilitarianism because the concept of maximising aggregate well-being is given priority over each individual. In other words Scanlon cannot accept the fact that society is organized around the idea of aggregate well-being instead of individual agreement⁵⁷. What is interesting about Scanlon's critique of utilitarianism, something that Scanlon's commentators have failed to see, is that it also applies to some aspects to Rawls's theory:

I have been criticising an argument for Average Utilitarianism that is generally associated with Harsanyi [...] But the objections I have raised apply as well against some features of Rawls's own argument. (Scanlon 1982, pp.123-4)

Thus we can see that Scanlon is not simply relying on Rawls's contractualism as a way of criticising utilitarianism, instead

⁵⁷ - To be precise, what Scanlon finds unacceptable in aggregate well-being is not so much the idea of individual well-being but the fact that this should be aggregative "One worthy feature of contractualist argument as I have presented it so far is that it is non-aggregative: what are compared are individual gains, losses and levels of welfare" (Scanlon 1982, p.123).

Scanlon is using his critique of utilitarianism as a way of distinguishing his own version of contractualism from Rawls's.

What Scanlon finds unacceptable in Rawls's contractualism is the fact that Rawls's theory runs the risk of subordinating moral motivations to an independently chosen social goal, namely, social cooperation as mutual advantage. To the extent that some idea of social cooperation is given priority over the original agreement, Scanlon rejects utilitarianism and Rawls's contractualism on the same grounds; in the case of utilitarianism the goal is maximizing aggregate well-being, while in the case of Rawls it is social benefit for mutual advantage.

After all, the concept of maximising aggregate well-being, central to utilitarianism, can be seen as a rationale for social cooperation, or in other words as the basic reason which determines the original agreement. As we have seen, what Scanlon finds unacceptable is not the concept of maximizing aggregate well-being *per se*, but the fact that this concept determines the original agreement (rather than being determined by it).

Thus according to Scanlon, both utilitarianism and Rawls's contractualism are upside-down: morality is instrumentally determined by an (independently determined) idea of social cooperation rather than allowing for social cooperation to be instrumentally determined by morality. It is to Scanlon's critique of Rawls's contractualism that I want to turn to next.

3.IV. Scanlon's Critique of Rawls.

Before launching on a detailed analysis of Scanlon's critique of Rawls, it ought to be emphasized that Scanlon's critiques are not meant to be a refutation of Rawls's theory, but an attempt to strengthen the latter. Indeed we can say that Rawls and Scanlon are engaged in the same project. This explains why there are many important similarities between Rawls's and Scanlon's contractualism. For example, the most important aspect of Scanlon's contractualism is that it rules out bargaining theory. According to Scanlon, bargaining is morally unacceptable since it results in:

being forced to accept an agreement by being in a weak bargaining position, for example because others are able to hold out longer and hence to insist on better terms. (Scanlon 1982, p.111)

Scanlon's rejection of bargaining theory is simply an echo of Rawls's views, in fact the function of a veil of ignorance in the original position is exactly that of erasing bargaining advantages⁵⁸.

Nevertheless there is also a major divergence between Rawls's contractualism and Scanlon's. Consider the following claim by Scanlon on the kind of contractualism he is prescribing:

⁵⁸ - See Rawls (1971), pp.139-140.

On this view (as contrasted with some others in which the notion of a contract is employed) what is fundamental to morality is the desire for reasonable agreement, not the pursuit of mutual advantage. (Scanlon 1982, p.115n)

It is not clear who Scanlon has in mind in the comment he makes in brackets; certainly it applies to Hobbes's social contract and his contemporary advocates, but as I have argued in Chapter 2 it may also apply to Rawls's theory. If the latter interpretation is correct we have a first indication of where Scanlon's contractualism differs from Rawls's - while mutual advantage plays an ambiguous role in Rawls's moral theory, it plays no part in Scanlon's moral set up.

In what follows, I will argue that Scanlon's rejection of mutual advantage as a moral concept indicates the gap between his view of contractualism and Rawls's in terms of the gap between the meta-ethical level and the ethical level. In other words while Rawls's contractualism operates at the ethical level, aiming to determine what principles are right or wrong, Scanlon's contractualism operates at a deeper level, a meta-ethical level, which aims to determine the nature of morality itself.

In order to see why Rawls's contractualism operates at the ethical level, I suggest we recall Rawls's idea of the original position. The original position is a decision procedure in ethics, that is to say, it is a device used to translate our moral intuitions into principles of justice. It follows that Rawls's

contractualism is silent about moral motivations, instead it is happy to assume our moral intuitions as raw data, and its function is to refine our intuitions and translate them into principles of justice. In doing so the original position upholds a specific view of egalitarianism. This can be defined as formal egalitarianism, to the extent that the principles chosen in the original position can be justified to each of their members.

We have seen that Rawls's contractualism is not concerned with our moral intuitions (i.e. with meta-ethical questions concerning the nature of morality), but with determining principles of justice (i.e. the ethical question of establishing what is right and wrong). Scanlon's contractualism is fundamentally different from Rawls's. In fact Scanlon understands that contractualism has an important role to play at the meta-ethical level, that is to say, the level of moral intuitions. The advantage of pitching contractualism at the meta-ethical level is that it enables Scanlon to theorise a concept of equality beyond formal egalitarianism.

In order to see the difference between Scanlon and Rawls, it is instructive to recall Scanlon's understanding of contractualism. According to Scanlon, hypothetical contracts:

include particular judgments as to what considerations should or should not be recognized as legitimate grounds on which members may refuse to accept given terms of cooperation.
(Scanlon 1977, p.51)

What is interesting to note is that in Scanlon's contractualism formal egalitarianism is replaced by a sense of equality based on what people could not reject as basis of agreement for terms of social cooperation.

It is because Scanlon's hypothetical contract is operating at the meta-ethical level rather than the ethical level that Scanlon's original agreement is different from what Rawls's envisioned, allowing for a different type of social cooperation compared to Rawls's.

Scanlon refers to the original agreement in his contractualism as 'reasonable agreement'. Here Scanlon finds no place for the veil of ignorance or rational choice theory in his model. The reasons for these omissions can be explained as follows: first, it permits Scanlon to avoid any correlation between his idea of an unanimous agreement and bargaining theory. Secondly, Scanlon rejects the claim that rational choice under uncertainty can model our moral intuitions in order to make these more vivid and precise.

In Scanlon's contractarian model, reasonable agreement is based on the moral motivation to act impartially, and impartiality is "the desire to be able to justify one's action to others on grounds they could not reasonably reject" (Scanlon 1982, p.116; emphasis added). It follows that according to Scanlon, the agreement is a direct consequence of this desire:

According to contractualism, moral argument concerns the possibility of agreement among persons who are all moved by

this desire, and moved by it to the same degree. (Scanlon 1982, p.111; emphasis added)

The difference between Scanlon's contractualism and Rawls's can be detected not only on the issue of the original agreement, but also on the issue of social cooperation. In fact just as Rawls's original position reflects the priority of mutual advantage within his conception of social cooperation, Scanlon's reasonable agreement upholds a specific idea of social cooperation, where mutual advantage is secondary: Scanlon's reasonable agreement indicates that social cooperation is determined first and foremost by the motive for impartial moral concern. According to Scanlon what is important to stress is the ethical dimension of social cooperation - that social cooperation among equals is to be valued for its own sake and not exclusively for its benefits.

The point is that according to Scanlon there is more to social cooperation than mutual advantage. As Scanlon points out: "membership in a cooperative association of equals may itself be counted an important and valuable good" (Scanlon 1977, p.58). In other words social cooperation among equals has an independent moral value separate from the benefits that can result from it.

If I read Scanlon correctly, principles of social justice ought to be grounded on the moral motivation to endorse impartiality, not on the pursuit of mutual advantage based on rational self-interest. This does not mean that social cooperation as mutual advantage plays no part in Scanlon's contractualism, but

simply that he subordinates the criterion of mutual advantage to the idea of agreement on a footing of equality.

I started this account of Scanlon's theory by claiming that Scanlon's contractualism deserves to be evaluated on its own right, and not as a parasitic byproduct of Rawls's contractualism. The distinction between Scanlon's contractualism and other theories is nowhere more evident than in the last few paragraphs of his article "Contractualism and Utilitarianism":

It is sometimes said that morality is a device for our mutual protection. According to contractualism, this view is partly true but in an important way incomplete ... The contrast may be put as follows. On one view, concern with protection is fundamental, and general agreement becomes relevant as a means or a necessary condition for securing this protection. On the other, contractualist view, the desire for protection is an important factor determining the content of morality because it determines what can reasonably be agreed to. But the idea of general agreement does not arise as a means of securing protection. It is, in a more fundamental sense, what morality is about. (Scanlon 1982, p.128)

I believe the above statement cleverly summarises the debate within the social contract tradition between the Hobbesian and Kantian camps, and it firmly establishes Scanlon as an authoritative advocate of the latter. By 'mutual protection'

Scanlon probably has in mind Hobbes's contractualist theory, where protection is the prime motivation behind the covenant, although there is no direct reference to Hobbes in Scanlon's article. If by 'mutual protection' we understand a specific case of 'mutual advantage', we have a first glimpse of where Scanlon's contractualism breaks away from other social contract theories. Although Scanlon doesn't reject the idea of social cooperation for mutual advantage tout court, he rejects the claim that mutual advantage is a determining factor in establishing the parameters of morality⁵⁹.

What makes Scanlon's contractualism different from Rawls, and in many ways better, is that Scanlon reassesses the balance between the two conceptions of social cooperation. That is to say, Scanlon subordinates social cooperation as mutual advantage to the idea of social cooperation on a footing of equality. Although both Rawls and Scanlon endorse a dual conception of social cooperation (as mutual advantage and as a morally desirable goal), Scanlon gives priority to the latter concept, while in Rawls's model the idea of mutual advantage plays an ambiguous moral role.

In conclusion, by pitching contractualism at the meta-ethical level, Scanlon is in a position to grasp a conception of equality that unquestionably rejects bargaining as a moral tool. Unfortunately the same cannot be said of Rawls's contractualism.

⁵⁹ – According to Scanlon the idea of social cooperation as mutual advantage has no moral force, instead social cooperation acquires a moral content only because it is subordinate to an agreement. In other words according to Scanlon what is moral in social cooperation is derived from the agreement, not vice versa.

Rawls's contractualism in general, and the original position in particular, re-introduces the logic of bargaining from the back door. This is why Scanlon claims that morality is not about mutual advantage, instead morality is about reasonable agreement.

By combining contractarian forms of argument with meta-ethical issues of moral motivation, Scanlon's theory is a valid and important contribution to the contemporary debate on social justice and liberal morality. The theory of reasonable agreement that I will offer in Chapter 4 is directly inspired by Scanlon's project. Nevertheless, it seems to me that there are some problems with Scanlon's theory that need to be resolved. It is to the limits of Scanlon's contractualism that I want to focus on next.

3.V. The Limits of Scanlon's Contractualism.

We saw earlier that the concepts of 'desire' and 'reasonableness' are of cardinal importance for Scanlon's contractualism, indeed they combine in the notion of contractualism as reasonable agreement. Unfortunately, these two concepts are also the weakest part in Scanlon's theory. There are two major problems with Scanlon's attempt to ground his moral theory on these two concepts: first, Scanlon is begging the question by defining the 'reasonable' in moral terms. Secondly, the psychological faculty of 'desire', central to Scanlon's argument, is insufficient (although necessary) to carry on its shoulders the weight of a moral theory.

In what follows I will consider these two fallacies.

A) The Fallacy of *Petitio Principii*.

I.M. Copi (1986, p.101) explains that the fallacy of *Petitio Principii*, or 'begging the question', is committed when one assumes as a premise for an argument the very conclusion it is intended to establish. I believe elements of this fallacy can be traced in Scanlon's argument. If we analyze Scanlon's idea of a non-utilitarian moral motivation, and his definition of reasonableness, we see that moral motivation centres on the notion of reasonableness, although reasonableness is in turn defined in terms of moral motivation.

The problem of circularity in Scanlon's argument has been stressed by Brian Barry:

According to [Scanlon's] alternative conception of morality, the primary motive for behaving morally for its own sake is simply "the desire to be able to justify one's actions to others on grounds they could not reasonably reject", where the basis for others' reasonable rejection of one's actions is given by their "desire to find principles which others similarly motivated could not reasonably reject". This may appear to be a circular definition of morality and the moral motive because the notion of 'reasonableness' already presupposes that people have some moral ideas. (Barry 1989a,

pp.423-4; emphasis in original)

I think Barry has pointed to a serious problem intrinsic to Scanlon's contractualism, although this does not invalidate Scanlon's theory tout court. In order to avoid this fallacy all that needs to be done is to define reasonableness in such a way that it does not encompass everything that is moral. There is no problem with acknowledging that the concept of reasonableness has a moral core, as long as this concept is not an identical reflection of morality. In other words we must avoid saying that to be reasonable is to act morally, where moral behaviour is reasonable behaviour.

I think it is possible to distinguish between a *thick* theory of moral motivation and a *thin* theory of moral motivation. The former is what Barry is objecting to, that is to say, an theory of moral motivation that entails the same meaning of morality that was supposed to be deduced from the account of moral motivation. On the other hand, a *thin* theory of moral motivation means that our motivations have a moral core, although this core *per se* is insufficient to explain the nature of morality unless it is combined with a moral and political context.

To return to Scanlon, what his theory of reasonable agreement requires in order to avoid this fallacy is a comprehensive definition of the notion of reasonableness. This definition must be *morally thin*, that is to say, it must reflect the moral aspect of an impartial agreement, while avoiding the risk of simply begging

the question. In other words, we must avoid saying that to be reasonable is to act morally, where moral behaviour is reasonable behaviour.

B) The Limits of 'Desire'.

The other weakness of Scanlon's model is its reliance on the idea of 'desire'. It seems to me that although 'desires' are a necessary ingredient in any contractualist model, they are not sufficient to carry the weight of a moral theory. I will illustrate this critique with the help of an analogy drawn from the literature on philosophy of action. In a famous article Donald Davidson (1963) explained rational behaviour in terms of two factors: belief and desires. He later modified his theory by adding a third factor, intentions, which stands between the belief-and-desire and the action⁶⁰.

Desires are a type of pro-attitude, that is to say attitudes directed toward action of a certain kind. Under pro-attitudes we find "desires, wantings, urges, promptings, and a variety of moral views, aesthetic principles, economic prejudices, social conventions, and public and private goals and values" (Davidson 1963, p.686). Obviously a desire or any pro-attitude per se cannot explain action; to will something is not the same as doing it,

⁶⁰ - D. Davidson (1985). See also C.J. Moya (1990).

although any rational action springs from a desire⁶¹. It follows that a desire must be followed by (a) a belief that some action will satisfy our desire, and (b) an intention, i.e. an all-out judgment, or positive evaluation of a way of acting. Thus Davidson concludes that we cannot ground our explanation of rational action exclusively on a desire or pro-attitude, instead 'desires' must be supported by 'beliefs' and 'intentions'.

I believe that the same medicine can be prescribed for Scanlon's reasonableness. Scanlon delineates reasonableness in terms of the "desire to find principles which others similarly motivated could not reasonably reject". While I agree with Scanlon that this 'desire' is necessary, it is not sufficient. The desire for reasonable agreement (i.e. Scanlon's moral motivation), or the desire to find adequately impartial principles, is unquestionably important, but it needs to be supported by beliefs and intentions: *beliefs* concerning the integrity of the desired end, for example the belief that reaching reasonable agreement is better than reaching other types of agreement or even no agreement at all, and *future intentions* concerning one's behaviour after the agreement has been reached, for example intentions to cooperate in the future even if one could be better off under a different set-up (i.e. the idea of fair-play)⁶².

⁶¹ – J. Elster even suggests that behaviour, to be rational "must stem from desires and beliefs that are themselves in some sense rational" (Elster 1985, p.62).

⁶² – It may be argued that when Scanlon defines reasonableness in terms of a 'desire', he is not referring to *first-order* desires (i.e. the desires or wants which shape a person's everyday actions

By comparing Scanlon's account of reasonableness with theories from philosophy of action my intention was only to reveal a weakness in Scanlon's theory, and point to possible ways of strengthening the latter. Unfortunately I have neither the space nor competence to take up this challenge in the present work. The point I am trying to make is that the two fallacies in Scanlon's account of reasonable agreement are testimony that this concept is still at its embryonic stage, and it will require much refinement before it can challenge utilitarianism as an alternative moral and political theory. Nevertheless I believe that Scanlon argument, even if not conclusive, is an invaluable contribution to the recent debate on contractualism and political theory.

So far I have argued that Scanlon's contractualism, grounded on the idea of reasonable agreement, is in many ways an improvement on Rawls's theory. Yet I have also shown that there are still some problems with Scanlon's theory, in particular the key concepts of the 'reasonable' and 'desire' are still problematic. In Part 3.VI I will consider two recent attempts to uncritically follow on Scanlon's footsteps. I will argue that both attempts are

and choices) but instead to second-order desires, where second-order desires incorporate beliefs and intentions. This distinction was first made by H.G. Frankfurt (1971), who argued that second-order desires, or desires about desires, will not necessarily coincide with first-order desires. I find Frankfurt's distinction to be more confusing than enlightening. If second-order desires are still desires, then my criticism of Scanlon's definition of reasonableness still stands. And if second-order desires are not desires, but only a catch-phrase for all sorts of motivations, then this category becomes too vague to be of any use while the analytical advantage of distinguishing between desires, beliefs and intentions is lost.

unsatisfactory, since the two respective theorists were not able to resolve the problems diagnosed in Scanlon's contractualism.

3.VI. Following in Scanlon's Footsteps.

The aim of this chapter was to vindicate the claim that contractualism is a necessary instrument for a moral theory grounded on impartiality. I argued that the *idea of agreement* (the willingness to compromise one's own conception of the good in order to allow room for others) is the key to the contractarian moment, indeed I argued that Scanlon's notion of a reasonable agreement is the first conscious attempt to build a contractualist theory around the *idea of agreement*, therefore raising contractualism to the meta-ethical level.

The theory of reasonable agreement is comprised of the consolidation of two separate concepts: agreement and reasonableness. So far the concept of agreement was given predominance, thus in the remaining pages of this chapter I want to briefly address the question of reasonableness⁶³.

Recently a number of political theorists have attempted to pursue Scanlon's project by developing the account of contractualism grounded on reasonable agreement. This is the case of Samuel Freeman and Thomas Nagel, who have explored the concept

⁶³ - The concept of reasonableness will be discussed in much greater detail in the following chapter: 'A Theory of Reasonable Agreement'.

of reasonableness in some detail. The reason why I feel it is important to look at the works of Freeman and Nagel is because their attempts to further Scanlon's project only highlights the two fallacies afflicting Scanlon's contractualism.

Since the publication of Scanlon's influential article on "Contractualism and Utilitarianism" in 1982, there has been a tendency to think that because contractualism applies to the meta-ethical level, indeed the agreement is based on the moral motivation to act impartially, then the concept of reasonableness must also be approached from a meta-ethical angle, or in other words the concept of reasonableness must also be explained in terms of such motivation.

Basically, this is how Freeman and Nagel have approached the concept of reasonableness. The problem with this approach is that it starts from the assumption that reasonableness is an agent-relative motivation, furthermore it assumes that reasonableness is a psychological rather than political concept. Because of these erroneous assumptions, the theories of Freeman and Nagel suffer from the same two fallacies as Scanlon's contractualism.

In what follows, I will argue that this approach is either futile (Freeman) or problematic (Nagel). It is futile because it only reiterates the problems inherent in Scanlon's contractualism; it doesn't help us to overcome them. On the other hand it is problematic because it steers us towards problematic questions that seem to be unanswerable.

Freeman (1991) claims that the primary significance of the

contractualism advocated by Scanlon lies in the way it relates motivation to practical reason (Freeman 1991, p.281). Indeed it is the specific account of moral motivation and its relation to practical reason that enables Scanlon to offer a philosophical account of morality⁶⁴.

In his detailed account of Scanlon's moral theory, Freeman makes two important points. First, that according to Scanlon's contractualism, there are no moral facts or properties prior to reasonable agreement. In other words, moral truth is defined by reference to the hypothetical agreement where free and informed persons could agree under ideal conditions. Secondly, that we have reasons to act in non-instrumental ways. Freeman tells us that contrary to the view of a growing number of moral philosophers (Foot, Williams, Harman, Gauthier), who argue that one can have reason to act morally only if it satisfies a moral desire or at least advances some other desire or non-moral interest (Freeman 1991, p.290), the moral motivations at the basis of contractualism are not for the sake of maximizing anything or effectively promoting a particular end, but for the sake of regulative, moral principles:

According to contractualism, moral motivation is ultimately principle-dependent; it rests on a desire to act on principles that could be justified to reasonable individuals. (Freeman

⁶⁴ - A philosophical account of morality entails (1) a foundational account of the subject matter of morality, and (2) a general moral view; see Freeman (1991), p.281.

1991, p.299)

In what follows I will attempt to show that Freeman's rejection of instrumental moral motivations, far from strengthening Scanlon's project, highlights the two fallacies I pointed out earlier.

The major problem with Freeman's account of contractualist morality is that he relies too heavily on psychological attributes, in particular the concept of reasonableness becomes a psychological characteristic. Freeman's rejection of instrumental moral motivations points to a moral view which risks being stripped of political reasons. Indeed it is not clear from his account of reasonable agreement why acting "for the sake of regulative (moral) principles" (Freeman 1991, p.293) should be preferred to instrumental reasons. It seems to me that Freeman's argument could stand on its feet only if certain principles are considered to be moral truths prior to reasonable agreement, although as we have seen this option is not open to Freeman.

The psychological attributes of Freeman's contractualism are no-where more evident than in his attempt to investigate the concept of reasonableness. According to Freeman, the concept of reasonableness is made up of two aspects: answerability and ideal codeliberation. Freeman tells us that:

individuals are answerable to one another for their conduct, the claims they make, and the expectations they have regarding

others' conduct and claims. (Freeman 1991, p.285)

Yet Freeman is keen to point out that answerability does not simply imply being held accountable, or responsible, for one's actions. Answerability "means that we ... be able publicly to justify our conduct, aims, and expectations on terms others could freely accept" (Freeman 1991, p.285).

Unfortunately Freeman does not tell us what these terms that others could freely accept are, hence his attempt at an explanation is as blurry as the concept he is trying to explain. Furthermore it appears that Freeman is as guilty as Scanlon of the petitio principii fallacy. Answerability is part of what it means to be reasonable, yet answerability is defined in almost identical terms to Scanlon's own definition of reasonableness: while Scanlon refers to reasonableness as the desire to find principles which others similarly motivated could not reasonably reject, Freeman talks of answerability as justifying one's conduct on terms others could freely accept.

Although answerability fails to elucidate the idea of the reasonableness, it would be unfair to dismiss Freeman's theory at this early stage. In fact answerability is only one part of what it means to be reasonable. The other part is a process of codeliberation. This process (which echoes Nagel's own solution) involves asking if my agent-centred deliberations satisfy moral requirements. What do these moral requirements amount to? Unfortunately Freeman's answer is always the same: "asking whether

my proposed actions, ends, and agent-centred reasons could be justified on grounds others could reasonably accept" (Freeman 1991, p.296).

Judging from Freeman's account of answerability and codeliberation, it is safe to conclude that he has been swimming in a circle, ending up where he had started out from. The idea of answerability and ideal codeliberation, central to Freeman's argument, cannot be considered a progress towards finding a valid conceptualization of reasonableness. It seems to me that if we want to make sense of the notion of reasonableness, we need to look beyond Scanlon's terminology, and look elsewhere for an explication of this key concept.

A similar critique to Freeman's account of reasonableness applies to Thomas Nagel. In 1987, Nagel made a first attempt to understand the idea of reasonableness by claiming that:

the standard of individual reasonableness is not merely a premoral rationality, but rather a form of reasoning that includes moral motives. (Nagel 1987, p.220)

I believe this is an important claim. Nagel understood that the notion of reasonableness must not be seen as antithetical to rationality, instead it is a form of rationality that entails moral motivation.

Four years later Nagel (1991) set out to investigate the concept of reasonableness in more detail, in the belief that this

concept may be instrumental in reconciling rationality with morality. Starting from the assumption that there is a conflict within each individual between the partiality of rational self-interest and the impartiality of moral motivations⁶⁵, Nagel envisioned the central problem facing moral and political philosophy in terms of discovering principles of conduct which accommodate both agent-relative and agent-neutral reasons for action. In other words Nagel wants to reconcile reasons specified by universal principles which nevertheless refer to features or circumstances of the agent for whom they are reasons, and reasons which depend on what everyone ought to value, independently of its relation to himself.

As I mentioned earlier, the concept of reasonableness is crucial for Nagel's enterprise, in fact Nagel's aim is to show how the concept of reasonableness can bridge the gap between the private sphere of rationality, and the public sphere of morality. Unfortunately Nagel is unable to solve the riddle he poses, instead he is forced to conclude that personal and impartial reasons cannot be reconciled in principles that are seen as acceptable from all points of view.

The inability to reconcile rational motivations with moral motivations leads Nagel to cast a pessimistic shadow on our future as well as on the likeliness of reasonableness to illuminate a way out of our misery. His conclusions are far from optimistic:

⁶⁵ - In the words of Nagel: "we are simultaneously partial to ourselves, impartial among everyone, and respectful of everyone else's partiality" (Nagel 1991, p.38).

At the moment I see no general solution to this problem. That is, there are, I suspect, no general principles governing both agent-relative, personal reasons and agent-neutral, impartial reasons, and their combination, which are acceptable from all points of view in light of their consequences under all realistically possible conditions. (Nagel 1991, pp.48-9)

While I agree with Nagel that the conflict within each one of us between the partiality of rational self-interest and the impartiality of moral motivations does exist, I am sceptical that this struggle can ever be resolved by appealing to the concept of reasonableness.

According to Nagel, reasonableness implies nothing more than standards of individual reasonableness, or in other words some kind of individual motivation to act morally. I believe this way of understanding reasonableness is the seed of Nagel's problems⁶⁶. Nagel's account of reasonableness is morally overloaded, with the result that the fallacy of petitio principii is once again present. As I pointed out earlier, while there is no problem with acknowledging that the concept of reasonableness has a moral core, the concept of reasonableness cannot be an identical reflection of morality.

⁶⁶ - Nagel may reply that his analysis is based on a valid deduction from Scanlon's own claims on impartiality. While not denying this, the fact remains that the limits of Scanlon's theory apply to Nagel's argument. It is also interesting to note the continuity in Nagel's thought from 1987 to 1991, in fact in the earlier article he writes that he is concerned with "the standard of individual reasonableness" (Nagel 1987, p.220; emphasis added).

It seems to me that by pitching reasonableness at the meta-ethical level, Nagel's approach is begging the question, in fact it is no great achievement to advocate reasonable agreement on the basis that agreement comes naturally once we all share the same commitment to act reasonably. Furthermore it makes reasonable agreement look trivial, since the problem which liberalism, and social justice in particular, is supposed to overcome is exactly that of a non-homogeneous social demands.

To recapitulate the argument, I have argued that contrary to the views of Freeman and Nagel, endorsing the view that contractualism applies to the meta-ethical level does not compel us to maintain that reasonableness is also a meta-ethical concept. It is a mistake to assume that reasonableness has to do with agent-relative motivations. It follows that the weakest part of Freeman's and Nagel's theories is their account of the concept of reasonableness, in fact their understanding of reasonableness focuses exclusively on agent-relative motivations.

In the following chapter, I will argue that if the concept of reasonableness is to be the grounding of reasonable agreement, and if reasonable agreement is going to overcome both (a) the *petitio principii* fallacy and (b) the over-reliance on the psychological faculty of desiring, then it is imperative to untie the knot that links the concept of reasonableness with agent-relative motivations. In order to make sense of the concept of reasonableness we must transcend the level of individual motivation, and look for a criterion of reasonableness which is

agent-neutral. Reasonableness must be considered a social and political concept, not a psychological motivation.

4. A THEORY OF REASONABLE AGREEMENT

In the previous two chapters, I critically evaluated the neo-contractualist theories of John Rawls and Thomas Scanlon. It should be clear by now that my intention was not to undermine their undertaking, but to point out weaknesses in their theories with the prospect of strengthening and thus pursuing the project they started. Following Rawls and Scanlon, my intention is to ground liberalism on a more adequate interpretation of egalitarianism compared to utilitarianism.

The character of the next two chapters will be more positive compared to the previous chapters. What I mean by this is that the point now is not of criticising, but of constructing, or in other words building upon the theories of Rawls and Scanlon in order to take neo-contractualism a step further.

The focus of this chapter will be on the concept of reasonableness, and how it relates to reasonable agreement. The reasonable agreement school of thought gets its name by advocating the consolidation of two separate concepts: 'Reasonable' and 'Agreement'. Yet while the latter term is relatively unproblematic, the idea of 'reasonableness' is still vague and unspecified, and although a growing number of scholars have backed reasonable

agreement as the central cog in the liberal engine, there is still very little agreement on what is considered reasonable or the meaning of reasonableness. This chapter will address this paradox inflicting reasonable agreement. Starting from the assumption that a clear and unambiguous understanding of the concept of reasonableness is a prerequisite if reasonable agreement is to be taken seriously (when is an agreement reasonable?), I will argue that the only sensible starting point must be an explication of the core term reasonableness.

This chapter is divided in six parts. Part 4.I takes issue against Scanlon's claim that the concept of reasonableness need not be substantively defined. Contrary to Scanlon, my impression is that a contractualist theory of reasonable agreement necessarily requires that the concept of reasonableness be specified. Indeed in Part 4.II, I will highlight the problems that derive from the failure to provide a precise definition of this concept.

In Part 4.III, I will argue that a prerequisite before attempting to define the concept of reasonableness is to disassociate this concept from the sphere of our moral intuition. Indeed reasonableness must not be synonymous with morality, furthermore it ought to be considered an agent-neutral concept.

Part 4.IV sets out the methodology for analysing and defining the concept of reasonableness. The way I intend to approach the elusive concept of reasonableness is by comparing it with the more common notion of rationality; this comparison will show that the reasonable and the rational are two separate concepts.

In Part 4.V, I will advance the proposition that the concept of reasonableness ought to be defined in terms of the maxim 'not asking for too much'. The advantage of this definition is that it turns reasonableness into an agent-neutral concept; reasonableness does not define our moral psychology, instead in the theory of reasonable agreement, the concept of reasonableness acts as a rationale which illuminates our moral intuitions. Part 4.VI will provide a short summary of the chapter.

4.I. Some Preliminary Remarks.

In the last five years, the idea that the concept of reasonableness ought to be defined has come under close scrutiny. For example, Thomas Scanlon (1992) and Peter de Marneffe (1990) have put forward arguments denouncing such undertaking. The starting point to our enquiry will be the sceptical views of Scanlon and de Marneffe concerning the project of defining the concept of reasonableness.

In recent years, moral theory has come under attack by two prominent philosophers, Bernard Williams (1985) and Michael Walzer (1983). Williams and Walzer see the aim of moral theory in terms of the discovery of objective truths about morality, for example, the view that there is a single correct theory of justice. To the extent that they consider this aim undesirable, Williams and Walzer are prepared to reject recent attempts to endorse and defend a liberal moral theory.

In a recent article, Scanlon (1992) rejects the idea that moral theory is about the search for clear and general principles which we can appeal to in order to decide which things are right or wrong. In order to defend moral theory from the charges of Williams and Walzer, Scanlon distinguishes between two aims of moral theory: Philosophical Enquiry, concerned with the nature of morality, and Moral Enquiry, concerned with establishing which actions or practices are right or wrong⁶⁷.

Scanlon is making an important point here. He wants to argue that a moral theory can still make valid statements about rightness and wrongness at the meta-ethical level, without necessarily having to endorse statable rules (or substantive moral principles) from which conclusions in particular cases can be derived. For example Scanlon views his own theory of contractualism as engaged in Philosophical Enquiry⁶⁸, whereas Harsanyi's utilitarianism is an example of a Moral Enquiry. Failure to distinguish between these two aims of moral theory leads to a narrow and incorrect view of moral theory, as the critiques by Williams and Walzer illustrate.

It is in order to subtract his contractualist theory from the critiques levelled at liberal moral theory by Williams and Walzer that Scanlon repudiates all attempts to scrutinize the concept of

⁶⁷ - In Chapter 3, I referred to Philosophical Enquiry as a meta-ethical theory, and Moral Enquiry as ethical theory.

⁶⁸ - See Scanlon (1992) p.5; Scanlon's formula that an act is wrong when it would be disallowed by any principles that no one could reasonably reject if they were seeking principles which could be the basis of informed, unforced general agreement, is not a substantive moral principle.

the reasonable. In other words, it is exactly because "a single unified set of substantive principles" is not waiting to be discovered that Scanlon undermines all attempts to explore the concept of reasonableness:

According to my version of contractualism, for example, we have reason not to be cruel to people or to break our promises to them because such actions would not be allowed by principles that they could not reasonably reject.....Even if this account is accepted, however, it does not follow that there need be a single unified set of substantive principles to which we can appeal to decide which things are wrong. I myself doubt that there are statable rules of this kind, from which conclusions in particular cases can be 'derived'. The grounds on which it is reasonable to object to cruelty, and unreasonable to reject a principle barring it, are different from the reasons why it would be unreasonable to reject a principle requiring the fulfilment of promises. (Scanlon 1992, pp.11-12)

What Scanlon is saying here is that we are deluding ourselves if we think that reasonableness can act as a common principle from which more specialized rules can be derived. In other words Scanlon is afraid that scrutinizing the concept of reasonableness in search for a substantive moral principle undermines the distinction between Philosophical Enquiry and Moral Enquiry, thus justifying

the views of Williams and Walzer. A similar argument to Scanlon's is put forward by Peter de Marneffe.

The starting point of de Marneffe's argument is given by the following question: can there be any values that are neutral between conceptions of the good? He argues that there is only one sense in which values can be neutral, namely, if "there are values that any reasonable person would accept as the basis of moral claims regardless of his or her particular conception of the good" (de Marneffe 1990, p.256).

What is of interest to us here is de Marneffe's treatment of the idea of a 'reasonable person'. Although he does not specify in any detail what reasonableness means, he is very explicit in specifying what reasonableness does not mean. In fact de Marneffe warns us against 'trivially' building into our concept of reasonableness the assumption that reasonable people will accept certain moral claims. He then adds:

What we can do, however, is point to certain values which are good candidates. If it is plausible to claim that any reasonable persons would accept a value as the basis of moral claims, then it is plausible to claim that this value is neutral in the relevant sense This is the project I take Rawls to be pursuing in presenting justice as fairness as a political conception of justice' (de Marneffe 1990, p.256).

de Marneffe is making a valid distinction between moral claims

and moral values. He correctly points out that while it is trivial to associate reasonableness with specific moral claims, what can be done is to formulate moral values that reasonable people would accept. Reasonableness entails the acceptance of certain values as the basis of moral claims, values that everyone ought to find acceptable; it does not entail the acceptance of specific claims.

I agree with Scanlon and Peter de Marneffe that the concept of reasonableness does not hide objective moral truths waiting to be discovered. The concept of reasonableness is not a modern day Platonic Form, nor is it the answer to all our moral dilemmas. Furthermore I agree with de Marneffe that if reasonableness is to gain substance and strength it is necessary to detach it from moral claims. Yet I also believe that if we are going to justify the contractualist idea of reasonable agreement, it is imperative to provide a tighter definition of reasonableness. In what follows I want to consider the risks the theory of reasonable agreement runs unless the concept of reasonableness is defined with precision.

4.II. Leaving Reasonableness Undetermined: Two Risks.

Not being able to provide a precise definition of the concept of reasonableness means that the idea of reasonable agreement is susceptible to the following two risks.

First, since at the moment the concept of reasonableness is, so to speak, up-for-grabs, we find that it is being adopted (and

abused) by a number of philosophers in order to add strength to their theories. This is the case with Thomas D. Perry (1976) and John Finnis (1980).

In Moral Reasoning and Truth Perry argues that it is possible to defend a theory of objective standards and evaluative truth, "based on the commonly accepted idea of reasonableness in moral reflection and argument." (Perry 1976, p.12). The core of Perry's argument is that reasonableness is a privileged value in ordinary moral discourse, whereby a reasonable judgment is both impartial and universal.

Perry seems to equate reasonableness with everything that is moral, indeed he claims that moral reasonableness cannot be coherently challenged or rejected by anyone if he or she is to engage in moral reasoning at all (Perry 1976, p.219). Yet Perry never attempts to define the concept of reasonableness, or tells us why we should accept his interpretation of reasonableness. It would appear as if under the guide of Perry the concept of reasonableness acquires a magical quality, since it can mean anything Perry wants it to while at the same time being unrefutable.

The problem with Perry's theory is that by appealing to the "commonly accepted criteria of reasonableness" Perry is begging the question. Contrary to Perry's optimistic views, it is not always obvious what is commonly accepted. Thus for example many liberals in the United States would argue that it is reasonable to be 'politically correct', while for many Germans in the 1920's it was reasonable to be anti-semitic. Perry tells us that reasonableness

is a moral concept, without telling us what it means.

A similar critique can be made to Finnis. Natural Law and Natural Rights is a long and complex book, where Finnis puts forward his theory of ethics as 'practical reasonableness'. By 'practical' Finnis means "with a view to decision and action" (Finnis 1980, p.12) and practical reasonableness is "reasonableness in deciding, in adopting commitments, in choosing and executing projects, and in general in acting" (Finnis 1980, p.12). The major problem with Finnis's argument is its circularity.

The circularity of Finnis argument becomes evident if we try to extrapolate a more precise definition of practical reasonableness from his work. I believe this definition lies in what Finnis calls the eighth requirement of practical reasonableness, namely, favouring and fostering the common good of one's community. This requirement is so important that it forms the basis of Finnis's definition of (general) justice:

Justice ... is in its general sense always a practical willingness to favour and foster the common good of one's communities, and the theory of justice is, in all its parts, the theory of what in outline is required for that common good. (Finnis 1980, p.165).

Once again we find that the concept of reasonableness is invoked to justify whatever the author thinks is right (in Finnis's case the common good of one's community) without any attempt to

explain why we should accept this particular interpretation of reasonableness.

The problem with Perry and Finnis is that they seem to embrace some idea of crude intuitionism, in fact their theories come across as nostalgic reiterations of the defunct natural law tradition. It is not surprising therefore that Perry and Finnis fail to justify their definitions of reasonableness, in fact in their moral theories the concept of reasonableness acts like a trump card: it wins the argument *a priori*.

The way in which the concept of reasonableness has been abused by Perry and Finnis justifies all the fears of Scanlon, which induced him to reject the view that the concept of reasonableness holds a substantive moral core. While I share Scanlon's preoccupations, my own response to this problem takes a different line from his: it is exactly in order to avoid more arbitrary abuses of the concept of reasonableness that it is necessary to give a tighter definition of reasonableness.

The second risk associated with the inability to provide a precise definition of reasonableness follows from the first risk, namely, it makes the contractualist idea of reasonable agreement vulnerable to charges of crude intuitionism.

The concept of reasonableness is at the heart of contractarian moral theory, hence understanding the former is a way to understand the latter. By the same token any problems associated with reasonableness are replicated on the contractualist theory of reasonable agreement; if the concept of reasonableness is treated

as a crude intuition, it follows that reasonable agreement too becomes engulfed in intuitionism.

Alan Hamlin (1989) argues that reasonableness and contractualism are intrinsically related. Echoing Scanlon's theory, Hamlin points out that any social arrangement or institution passes the contractarian test if no one can reasonably reject it. He then continues: "Of course, all the work here is being done by the requirement of 'reasonableness'". (Hamlin 1989, p.98). According to Hamlin, the requirement of reasonableness is what contractualism is all about. Indeed discussing the theme of contractarian rights, Hamlin points out that reasonableness is the core of contractualism:

Again, this formulation loads all of the work onto the requirement of 'reasonableness', but this is precisely the intent of any contractarian. Given a particular notion of what is individually reasonable, the contractarian method is the means of building from that notion into principles to govern social interaction. (Hamlin 1989, p.99)

If we accept Hamlin's view of contractualism as essentially correct, it follows that a contractualist theory of reasonable agreement cannot avoid tackling the problem of investigating the concept of reasonableness.

Leaving the concept of reasonableness undetermined has led some to feel that there is nothing more to this concept than a mere

moral intuition, which means that the contractarian theory of reasonable agreement is by implication nothing more than a sophisticated moral intuition dressed up as a constructivist theory.

The charge of intuitionism has been made forcefully by Allan Gibbard in his review article of Brian Barry's Theories of Justice. Gibbard (1991) argues that if the concept of reasonableness is given such prominence in a theory of justice, as in the case of Barry's theory⁶⁹, then it is unacceptable to depend for an elucidation of the term 'reasonableness' on intuitive understandings. It is unacceptable because it reduces reasonableness to brute intuitions, and intuitions are at best a weak and controversial foundation for a theory of justice⁷⁰. In the words of Gibbard:

If, though, the unreasonableness of rejecting certain principles is just a brute moral fact, whispered to us by the voice of intuition, then I worry. Argument stops somewhere, to

⁶⁹ - "Barry's test is to ask which agreements agents of certain kinds could reject and still be reasonable. This puts the burden on the moral term 'reasonableness'". (Gibbard 1991, pp.278-9).

⁷⁰ - In the previous chapter I mentioned that Nielsen accuses Kymlicka of intuitionism. Nielsen's argument can be compared to Gibbard's: "But if there are any accounts that are by now widely recognized to be non-starters, it is intuitionism and natural law theories where we in some mysterious way must just have direct access to the truth - indeed, even on some accounts, the certain truth - of certain moral propositions. How then does Kymlicka know, and how can we know, that his fundamental substantive moral claims, claims not subject to agreement, are true or justified?" (Nielsen 1992, p.94).

be sure, but I hope it does not have to stop with brute intuitions of what is reasonable and what is not. If someone tells me a demand is unreasonable, I want to ask why. (Gibbard 1991, p.279)

I believe Gibbard's challenge is important, and ought to be taken seriously. The accusation of intuitionism deserves a response. If reasonable agreement is to be the foundation of a comprehensive liberal political theory, then the concept of reasonableness cannot be left undefined.

To recapitulate the argument so far, I have argued that failure to provide an accurate definition to this concept leads to two unwanted consequences. First, it facilitates the seizure of this concept by unscrupulous philosophers who preach the virtue of reasonableness in an attempt to defend their subjective moral views. Indeed the theories of Perry and Finnis are confirmation that this has already been the case, with the result that the concept of reasonableness has lost any credibility. Secondly, failure to define the concept of reasonableness means that the contractualist idea of reasonable agreement is vulnerable to the charge of crude intuitionism.

In what follows, I will argue that in order to avoid these two traps, it is imperative to give an accurate definition to the concept of reasonableness, furthermore it is fundamental that this concept does not become a catch-all phrase for all that is morally good.

4.III. Defining Reasonableness.

So far I have argued that it is a grave risk to leave the concept of reasonableness undefined. At the same time one has to be careful not to define this concept in such a way that will undermine Scanlon's important distinction between Philosophical Enquiry and Moral Enquiry.

In searching for a definition of the concept of reasonableness, it is important to always remember that the foundation of a contractualist moral and political theory (i.e. the foundation of reasonable agreement) is in the moment of hypothetical agreement, based on impartial moral motivation. The concept of reasonableness should not replace the moment of agreement, and moral motivation in particular, as the focus of moral theory. Instead defining the concept of reasonableness must be seen as complementary to moral motivations, or in other words as explaining and justifying our impartial moral motivations.

The problems with defining the concept of reasonableness that Scanlon alludes to only arise if we treat this concept as integral to our moral motivations. Yet the brief accounts of Freeman and Nagel in previous chapter demonstrate that there is a tendency to assume that the concept of reasonableness is part of our moral motivations. This is exactly where the problem lies. If we start from the assumption that reasonableness is part of our moral motivations, then we are left with an insurmountable problem,

furthermore any attempt to define the concept of reasonableness substantiates Scanlon's premonitions.

Seeking a definition of reasonableness, based on the assumption that this concept reflects our moral psychology, leads to a dead end. Once we accept that reasonableness is a moral motivation, we are vulnerable to Barry's reproach that the argument is guilty of circular logic⁷¹. To recycle one of Barry's aphorisms, the project of defining reasonableness in moral terms and then deducing normative principles from the idea of reasonable agreement, can be compared to a conjurer putting a rabbit in a hat, taking it out again and expecting a round of applause.

Yet I see no reason why we must accept the assumption that identifies reasonableness with our moral motivations. A theory of reasonable agreement does not require that reasonableness be considered a moral motivation, since in a contractualist theory the moral motivation comes from the idea of agreement, where the idea of agreement is defined in terms of the willingness to compromise one's conception of the good in order to make room for others.

But if reasonableness is not a moral motivation, what is it? Why do we appeal to an agreement that is reasonable? The brief answer is that the concept of reasonable enables us to bridge the gap between the meta-ethical level of agreement (i.e. our moral intuitions), and the political sphere, where institutions are evaluated on normative grounds.

⁷¹ - For an account of the problem of circularity, see Chapter 3 above.

In order to fully understand the idea that reasonableness can be divorced from the sphere of moral motivations, it will be instructive to analyze Gibbard's critical assessment of reasonable agreement. According to Gibbard (1991), in order to avoid the trap of crude intuitionism, we must be able to display a convincing rationale for our intuitions, i.e. a rationale that gives us a kind of knowledge of what is reasonable. Obviously Gibbard assumes that reasonableness is a moral intuition, hence we need a rationale in order to explain what is reasonable and what is not. In the words of Gibbard:

We do have intuitions, I agree, about what is reasonable and what is not. The question is how we should regard these intuitions. Perhaps they can be given an illuminating rationale. Then we can see the intuitions as implicitly responding to this rationale. If the rationale is good, we can regard the intuitions as giving us a kind of knowledge of what is reasonable. (Gibbard 1991, p.279)

I believe that Gibbard is inadvertently providing us with the answer to all our problems; all we need to do is put Gibbard's model up-side-down. The point is that we must start from the assumption that reasonableness is not our moral intuition, instead our intuitions are given by (a) the idea of agreement, and (b) the desire to justify our actions to others. If we accept this, then we can see that the concept of reasonableness provides (to use

Gibbard's terminology) the 'rationale' that illuminates our intuitions.

It seems to me that the above argument helps us to understand both the concept of reasonableness and the notion of reasonable agreement: *the concept of reasonableness is the rationale which qualifies our moral intuitions, where the idea of agreement is the basis of our moral intuitions.* In the theory of reasonable agreement, the concept of reasonableness qualifies the agreement, or in other words reasonableness is the rationale that illuminates our moral intuitions.

An important aspect of the above argument is that it placates Scanlon's worries concerning the distinction between Philosophical Enquiry and Moral Enquiry while still allowing us to tighten the definition of reasonableness. Seeking a definition of reasonableness need not imply that this concept acts as a common principle applicable to all issues, from cruelty to keeping promises, as Scanlon fears. Instead as Barry points out, what we need to do is to specify further the notion of reasonableness, not substitute some simpler and more tractable notion⁷². The only way to do this is by defining the concept of reasonableness without appealing to moral motivations.

In Part 4.IV, I will argue that the first step towards defining the concept of reasonableness is to differentiate it from the better known (and hegemonic) idea of instrumental rationality, while in Part 4.V I will attempt to give a precise definition of

⁷² - See Barry (1989a), p.426.

reasonableness.

4.IV. Reasonableness vs Rationality.

It is important to recognize that the concept of reasonableness deserves to be analyzed as a distinct and original concept, with full autonomy from other moral, social and political terms. For example, the concept of instrumental rationality has always subsumed the concept of reasonableness under its hegemony. In what follows I intend to show that the uniqueness of reasonableness can be best appreciated when contrasted against the concept of rationality. By referring to the writings of Chaim Perelman and John Rawls, I will argue that the concepts of the reasonable and reasonableness⁷³ are conceptually distinct from the idea of rationality; justifying this claim is crucial in order to show that a reasonable agreement has nothing in common with what can be called a rational agreement⁷⁴. But first it is useful to say something on the etymology of the terms 'rationality' and 'reasonableness'.

In its everyday use, the term 'reasonable' is used as a synonym for 'rationality'. Stemming from the same latin root

⁷³ - In this chapter, and throughout the thesis, I will use the terms 'the reasonable' and 'reasonableness' interchangeably.

⁷⁴ - A rational agreement is an agreement where the parties are moved by a rational desire to maximise their ends, irrespective of others; Rawls's original position and Gauthier's bargaining-based agreements would fall under this category.

(RATIO-), it is not surprising that in the past the 'rational' and the 'reasonable' were used as synonyms. Indeed the Oxford English Dictionary (hereafter OED) shows that there is a strong affinity between these two terms: to be 'rational' means "Having the faculty of reasoning; endowed with reason", while 'reasonable' also implies being "Endowed with reason. = RATIONAL".

Although the etymology of the words apparently fails to distinguish the reasonable from the rational, it cannot be denied that in the last fifteen years the way these terms have been referred to by some political philosophers points to a degree of rupture in the previous synonymous relationship. In what follows I will take the first steps towards distinguishing the concept of the reasonable from that of rationality. I will first examine two recent attempts, by Perelman and Rawls, to distinguish the reasonable from the rational. I will then draw some speculative conclusions on the conceptual fracture between the reasonable and the rational, claiming that only the reasonable is a normative concept.

The conceptual distinction between the reasonable and the rational has troubled legal theorists for a long time⁷⁵. It is not surprising therefore that in recent years, one of the most explicit attempt to scrutinize the concept of reasonableness and dissect it from that of rationality comes from a legal theorist with a strong background in formal logic and highly respected in political

⁷⁵ - See Aarnio (1987). For an account of the 'Reasonable Man' in jurisprudence, see MacKinnon (1990).

theory: Chaim Perelman.

In his provocative article "The Rational and the Reasonable"⁷⁶, which has sadly been neglected in recent debates, Perelman argues that reasonableness and rationality function on two different wavelengths; rationality is instrumental, being concerned with the means to ensure predetermined ends, while reasonableness is normative, dealing simultaneously with both means and ends.

Perelman claims that rational behaviour, being concerned only with each individual's calculation of the best means towards an end, is stripped of the social dimension. On the other hand reasonable behaviour is characterized by a strong social motivation which Perelman traces into the search for acceptability. Thus contrary to rational actions, reasonable actions are acceptable to everyone, not just to the person who is doing the action:

'Rational' man separates reason from the other human faculties and shows a unilateral being functioning as a mechanism, deprived of humanity and insensible to the reactions of the milieu: he is the opposite of the reasonable man [Reasonable man] is guided by the search, in all domains, for what is acceptable in his milieu and beyond it, for what should be acceptable by all. Putting himself in the place of

⁷⁶ - Lecture delivered at the International Symposium 'Rationality Today' held at the University of Ottawa in October 1977. Proceedings published by the University of Ottawa Press 1979, 213-224. Also reprinted in Chaim Perelman (1979) The New Rhetoric and the Humanities. All references are taken from the 1979 publication.

others he does not consider himself an exception but seeks to conform to principles of action which are acceptable to everyone. (Perelman 1979, p.118).

One of the most appealing characteristics of Perelman's distinction between the rational and the reasonable is that he defines the latter appealing to two Kantian ideas: the categorical imperative, which makes the universal the criterion of morality, and the Golden Rule. This invocation of Kant is the bridge connecting Perelman's views to those of Rawls.

According to Rawls, we ought to make a distinction between rationality and the reasonable, where the latter should have priority over the former. In his 1980 Dewey Lectures Rawls argues that the concept of reasonableness is autonomous and independent from the notion of rationality, and ought to be defined on different grounds than rationality: the reasonable expresses a conception of the fair terms of cooperation, while the rational expresses a conception of each participant's rational advantage, in other words what as individuals each participant is trying to advance. (Rawls 1980, p.528).

It follows that according to Rawls, the reasonable has a moral underpinning⁷⁷, making reasonableness and fairness closely related. Furthermore the reasonable and the rational do not stand on equal rank, instead Rawls clearly believes in a lexical priority between

⁷⁷ - "Reasonable persons, that is, ... persons who have realized their two moral powers ..." (Rawls 1989, p.236).

these two concepts:

[T]he Reasonable presupposes and subordinates the Rational. It [the Reasonable] defines the fair terms of cooperation acceptable to all within some group of separately identifiable persons, each of whom possesses and can exercise the two moral powers ... The Reasonable subordinates the Rational because its principles limit, and in a Kantian doctrine limit absolutely, the final ends that can be pursued. (Rawls 1980, p.530)

By distinguishing the reasonable from the rational, and claiming that the former limits the latter, Rawls is trying to find a place for both the reasonable and the rational in his theory of justice. This is the way Rawls's controversial claim of a "social cooperation for the mutual benefit of all" (Rawls 1982, p.164) should be interpreted; by social cooperation Rawls implies "fair terms of cooperation", or reasonable terms of cooperation, whereas mutual benefit presupposes rationality.

In his latest work Political Liberalism, Rawls attempts to specify more clearly the difference between the reasonable and the rational. He claims that rational agents lack the moral sensibility that underlies the desire to engage in fair cooperation (Rawls 1993a, p.51), or in other words that they lack a sense of justice and fail to recognize the independent validity of the claims of others (Rawls 1993a, p.52). On the other hand reasonable persons

desire a social world in which they, as free and equal, can cooperate with others on terms all can accept (Rawls 1993a, p.50).

In his latest work Rawls is also more specific about what he understands by the concept of reasonableness. He claims that reasonableness has two aspects: (a) the willingness to propose and honour fair terms of cooperation among equals, and (b) the willingness to recognize the burdens of judgment - the sources of reasonable disagreement - and to accept their consequences. Rawls acknowledges that both aspects of the reasonable are closely connected with Scanlon's principle of moral motivation (Rawls 1993a, p.49n), indeed Rawls is keen to show that Scanlon's principle is more than a psychological principle of motivation. In §7 and §8 of Lecture II: The Power of Citizens and Their Representation⁷⁸, Rawls argues that although the basis of moral motivation is partly psychological, it is not exclusively psychological since it concerns the fundamental question why anyone should care about morality at all. In the words of Rawls, moral psychology is philosophical, not psychological.

While Rawls's analysis of the reasonable and its difference from rationality is a vast improvement on the vague assertions made by Perelman, Rawls's latest account raises some new problems which he is unable to overcome: Rawls's account of reasonableness is nothing more than a theory of moral motivation, and as such it

⁷⁸ - While Lecture 2 is taken from material already published in 1980, §7 and §8 represent an original contribution by Rawls and the material discussed here did not appear in the three Dewey lectures on "Kantian Constructivism in Moral Theory".

undermines the reason for constructivism.

Rawls distinguishes between three types of desires: object-dependent desires (where the object of desire is divorced from moral conceptions, for example a desire for sex, food or shelter); principle-dependent desires (where the object of desire is given by certain rational or reasonable principles, such as the principles of utility and lexicographical preferences for rational principles, and principles of fairness and justice for reasonable principles); and finally conception-dependent desires. Conception-dependent desires are according to Rawls the most important of the three, therefore require special attention.

Conception-dependent desires are principle-dependent desires with a difference, namely, the principles must be related to a political ideal. In Rawls's case, the political ideal is that of citizenship as characterized in justice as fairness. Rawls's debt to Scanlon here is obvious⁷⁹, in fact it is by emulating Scanlon's idea of moral motivation that Rawls elaborates a moral psychology of the person: Rawls claims that citizens have a "reasonable moral psychology" (Rawls 1993a, p.86), and the salient feature of Rawls's sketch of the moral psychology of the person is that it is philosophical, not psychological. In the words of Rawls:

I stress that it is a moral psychology drawn from the political conception of justice as fairness. It is not a

⁷⁹ - Not surprisingly Rawls claims that Scanlon's moral motivation is a conception-dependent desire; see Rawls (1993a), p.85n.

psychology originating in the science of human nature but rather a scheme of concepts and principles for expressing a certain political conception of the person and an ideal of citizenship. (Rawls 1993, pp.86-7).

It seems to me that Rawls's account of moral motivation (reasonable moral psychology) assumes the very conclusions that he wants to defend (an ideal of citizenship). The problem with Rawls's account is that constructivism becomes redundant. That is to say, if everyone is moved by the kind of moral motivation Rawls is describing, then there would be no fundamental disagreement and contractualist would not be doing any work⁸⁰.

The conception of reasonableness that I want to defend differs from Rawls's conception on the issue of the very nature of reasonableness. According to Rawls's account, reasonableness is a motivation intrinsic to the beholder of reasonable moral psychology, while according to my conception, reasonableness is something that can be recognized from outside, that is to say, reasonableness is an agent-neutral criterion, which in the first place is external to the psychology of the person.

While I share the view of Perelman and Rawls that rationality and reasonableness are two distinct concepts, the distinction is not that rationality lacks a moral dimension while reasonableness

⁸⁰ - In Chapter 3 I argued that Scanlon's reasoning is dangerously circular, and that he commits the fallacy of petitio principii. I believe the same criticism applies to Rawls's account of reasonableness and reasonable moral motivation in Political Liberalism.

is a moral term. In fact I believe that reasonableness shares with rationality some important characteristics, for example their criterion is independent of moral desires.

An example will make this point clear. A rational act is one that can be recognized as such by everyone, notwithstanding their views. For example if I am in New York and I have a desire to get to Los Angeles within the next 12 hours or something catastrophic will happen (let's imagine that hospital authorities require my signature before operating on my wife, and she is not going to live for more than 12 hours), and I have the choice between a direct flight, a three-day train journey or a 6-day drive, it is rational for me to go to the airport. What is rational about my decision is not so much the fact that I have a desire to save my wife's life, but that I want to get from A to B in the shortest possible time⁸¹.

I believe that a reasonable act must also be universally recognized by everyone, notwithstanding the differences between persons. The criterion of reasonableness must be independent of our desires, that is to say, what is considered 'reasonable' cannot be intrinsic to the moral psychology of the person⁸².

The distinction between rationality and reasonableness is the following: rationality is essentially an individualistic virtue or

⁸¹ - The same reasoning would apply if I had a desire to see a baseball game in California. The reason for the trip does not interfere with the rationality of my decision concerning the mode of transportation.

⁸² - In Part 4.V below I will argue that the maxim 'not asking for too much' gives a definition of reasonableness which is not reducible to the object of our desires.

quality, while reasonableness is essentially a social virtue or quality. Being reasonable and being rational both share the idea of being sensible⁸³, yet while rationality is increasingly becoming associated with prudence, reasonableness is becoming associated with social virtue. As Barry points out, reasonableness "is now seen as a more social virtue than rationality" (Barry 1987, p.420).

The social quality of reasonableness, as opposed to rationality, was appreciated by W.M. Sibley more than 40 years ago:

To be reasonable ... is to see the matter - as we commonly put it - from the other person's point of view, to discover how each will be affected by the possible alternative actions; and, moreover, not merely to 'see' this (for any merely prudential person would do as much) but also to be prepared to be disinterestedly *influenced*, in reaching a decision, by the estimate of these possible results. (Sibley 1953, p.557)

It is interesting to note that Rawls models his conception of reasonableness on Sibley's definition, indeed in Political Liberalism Rawls points out that his discussion of reasonableness accords with Sibley's basic distinction between the reasonable and the rational (Rawls 1993, 49n).

That Rawls was taken by Sibley's definition is not surprising,

⁸³ - The OED tells us that to be reasonable means "having sound judgment; sensible, sane", while being rational implies being "agreeable to reason; reasonable, sensible; not foolish, absurd, or extravagant".

in fact there are strong affinities between Scanlon's account of moral motivation, being a desire to justify our actions to others on grounds they could not reasonably reject, and the following definition of reasonableness by Sibley:

I must justify my conduct in terms of some principle capable of being appealed to by all parties concerned, some principle from which we can reason in common. (Sibley 1953, p.557)

Of course Sibley saw reasonableness as a moral virtue, as essentially related to the disposition to act morally, and Rawls follows in Sibley's footsteps to read into the concept of reasonableness a moral motivation. I think Sibley and Rawls are both wrong to identify the concept of reasonableness with our moral motivations⁸⁴, although Sibley was right to emphasize the social quality of reasonableness as the distinctive factor of this concept compared to rationality⁸⁵.

I am not denying that the concept of reasonableness has a

⁸⁴ - In Chapter 3 I mentioned Barry's warning that it is not sufficient to substitute reasonableness with some simpler and more tractable notion, instead what we need to do is to specify further the notion of reasonableness. It seems to me that Sibley does exactly what Barry says we should not do: "Reasonableness thus requires impartiality, 'objectivity'; it expresses itself in the notion of equity." (Sibley 1953, pp.557-558).

⁸⁵ - Rawls also accepts that rationality is an individual virtue while reasonableness is a social virtue, in fact echoing an argument by Freeman (1990, p.141-147), Rawls claims that "A further basic difference between the reasonable and the rational is that the reasonable is public in a way that the rational is not" (Rawls 1993a, p.53). Yet in Rawls the social virtue collapses in moral power; see Rawls (1993a), pp.48-54.

moral dimension. After all, if we endorse the argument that the distinction between rationality and reasonableness is that rationality is essentially an individualistic virtue or quality, while reasonableness is essentially a social virtue, then it becomes impossible to strip the concept of reasonableness of all moral connotations (that is, if by morality we understand a certain type of social behaviour, where my actions interact with or have consequences on others). The affinity between moral concepts and social circumstances explains why Sibley gives reasonableness a moral coating⁸⁶.

Nevertheless, the concept of reasonableness is not an identical reflection of the concept of morality, or in other words reasonableness cannot be a synonym for morality. In what follows, I will argue that the concept of reasonableness is an agent-neutral concept which has the function of illuminating the moral intuitions of impartiality and the *idea of agreement*.

4.V. Rethinking Reasonableness.

So far I have tried to show that there is some evidence of a conceptual fracture between these two concepts. But what exactly does being reasonable or reasonableness amount to? How does the

⁸⁶ – I am grateful to Margaret Moore for a discussion on the morality of reasonableness.

reasonable relate to reasonable agreement? In this section I will attempt to answer these questions.

I believe Barry (1987) comes closest to pin-pointing the social quality of reasonableness when he reminds us that according to the OED, reasonableness may also mean 'not asking for too much'. In what follows, I will start by exploring the idea of 'not asking for too much', followed by an attempt to relate the concept of reasonableness, as defined above, to the idea of an agreement: hence reasonable agreement.

The intuitive idea of reasonableness as 'not asking for too much' needs to be examined and qualified further if it is to prove useful to the contractualist theory of reasonable agreement. There is an explicit and an implicit side to the idea of 'not asking for too much'. Explicitly the notion of 'too much' points to a quantitative measurement, while implicitly it implies a comparative relation.

Thus in ordinary speech when we say that something is 'too much' we have in mind a quantitative relationship whereby something is *too much of something-else* which is in limited supply. This first qualification finds support in the OED, where under 'Reasonable' we find the following explication: "Of such an amount, size, number, etc., as is judged to be appropriate or suitable to the circumstances or purpose".

Implicitly, 'not asking for too much' assumes that one's claim is *too much compared to some other claim for the same limited good*, hence its intrinsic comparative quality. Once again this second

qualification finds support in the OED, where it is said that the term 'Reasonable' may also be used to mean 'Proportionate'.

The important characteristic of the notion of something being 'too much' is that this notion assumes the scarcity of some good or resource. Thus in order to establish if a demand is 'too much', we must first have an idea of what the subject of the demand is. All demands are demands *for something that is usually in limited supply*. It is that *something* which gives the notion of 'not asking for too much' the social dimension, since all demands are evaluated within the context of scarcity of resources, and everyone⁸⁷ has an equal right to make a claim to these resources.

If we take the idea of 'not asking for too much' as a whole, that is to say combining its explicit and implicit connotations, together with the awareness of a context of scarcity in which demands are made, we get an approximate idea of how reasonableness can be defined.

A demand is reasonable if it does not ask for too much of a scarce resource compared to other demands for the same resource.

The above definition of reasonableness is important because it puts our moral motivations (defined by the idea of agreement) within the appropriate social and political context. In other words the concept of reasonableness acts as the rationale that

⁸⁷ - Everyone here applies also to those who are not yet born.

illuminates our intuitions. It follows that intuitionism plays a necessary role in contractualism, but not a sufficient role.

Brian Barry (1989b, pp.271-282) points out that there are three possible views about the relation between constructivism and intuitionism: that constructivism and intuitionism are entirely independent; that constructivism is simply a variety of intuitionism; and that constructivism cannot be entirely independent of intuitionism but is not entirely reducible to it either⁸⁸.

Of the three views, Barry defends the third alternative, claiming that "constructivism [is] dependent on but not reducible to intuitionism" (Barry 1989b, p.282). I believe that defining the notion of reasonableness in terms of the maxim 'not asking for too much' is crucial if contractualism is to maintain a link with intuitionism without being entirely subordinated to it; if reasonableness refers to a social and political context of scarcity, then the notion of reasonable agreement cannot be reduced to intuitionism.

The time has come to reconcile the above definition of reasonableness with the contractarian argument. In other words to reconcile the idea of the reasonable, with the moment of agreement.

⁸⁸ - Logically speaking, there is also a fourth alternative, namely, that intuitionism is simply a variety of constructivism. I think this fourth alternative is a non-starter. Perhaps the best attempt to provide a constructivist account of our motivations is found in Gauthier's Morals by Agreement, although it may be argued that Gauthier is unable to avoid moral intuitions. For a criticism of Gauthier along those lines, see Goodin (1993) and Lehning (1993).

In Chapter 3 I pointed out that the notion of agreement is at the heart of Scanlon's model, an agreement to cooperate on terms that everyone could reasonably accept (an agreement people could make under circumstances where no one was ignorant or coerced).

While the idea of agreement defines our moral intuition, crude intuitionism cannot be the foundation for a political theory. For a moral theory to become a political theory, it is fundamental to bridge the gap between our moral intuitions and the political sphere; this is the role performed by the concept of reasonableness, and the maxim 'not asking for too much'.

The question I want to address in this section is how the definition of reasonableness in terms of 'not asking for too much' fits in with Scanlon's contractualism. As I said before, the contribution of Scanlon's moral theory is to have offered a contractualist account of moral motivation. Scanlon says that according to contractualism, the source of motivation is the desire to be able to justify one's actions to others on grounds they could not reasonably reject.

It is important to see that Scanlon is not simply appealing to an inter-subjective equilibrium, where everyone agrees in virtue of giving their consent. In other words Scanlon is not simply saying that we have a desire that others do not reject our actions, for this would be nothing more than justifying an agreement based on the consensus of the people. The problem here is that such equilibrium does not have moral force, for example a person may be spontaneously willing to give up her freedom and become a slave,

thereby creating a consensus agreement between the slave and her owner, yet this equilibrium *per se* does not have moral force.

I believe there is an objective standard at the core of Scanlon's formula, albeit hidden, which qualifies our moral motivations. Indeed a closer examination of Scanlon's formula reveals that Scanlon is appealing to certain *grounds* that others could not reject. It is these grounds that are the basis for the equilibrium between my actions and the refutability of others. Having established this fact, the question I want to address now is the following: what are the *grounds* that others could not reject?

I believe that establishing these grounds is the primary function of the concept of reasonableness. More specifically, I believe that the maxim 'not asking for too much' identifies the grounds that establishes what others could not reasonably reject. In other words the maxim 'not asking for too much' sets the ground for determining if certain actions can be rejected by others or not. Although the category of others in Scanlon's contract does not include everyone but only those moved by a similar desire to reach an agreement (in other words Scanlon does not have to justify his actions to everyone but only to others similarly motivated), we need a standard for determining the validity of our motivations. I believe the idea of 'not asking for too much' can shed some light on this requirement.

It follows that the definition of reasonableness as 'not asking for too much' can be used as the key to understand the idea of reasonable agreement. Indeed I want to argue that reasonable

agreement is a desire to find an agreement under circumstances of scarcity, where the agreement itself cannot be divorced by the fact that everyone who is making a claim for resources is thereby excluding others.

If reasonableness involves 'not asking for too much' it follows that reasonable agreement is an agreement reached when (a) everyone is moved by a similar moral motivation, and (b) no-one is asking for too much, since not asking for too much is the only demand that others could not reasonably reject: *an agreement is reasonable when no-one seeking agreement is asking for too much.*

I conclude therefore that reasonableness concerns fair demands for scarce resources, and that reasonable agreement is an agreement concerning the use or distribution of scarce resources that takes account of the fair demands of others⁸⁹.

4.VI. Conclusion.

⁸⁹ - Of course one may object that the maxim 'not asking for too much' does not ensure that others are not made worse off, since it may well be the case that even asking for very little of a scarce resource may make others infinitesimally worse off. For example my demand for better public lighting in my street in North London may cost the state very little but still make someone in Liverpool worse off by a fraction of a penny. While this objection cannot be refuted, it applies only if the maxim 'not asking for too much' is seen as a moral virtue which we can appeal to determine what is right and wrong. But this is not how I intend to use the maxim, hence the objection is out of place. The maxim 'not asking for too much' is not a rule from which conclusions in particular cases can be derived (this is Scanlon's objection to a normative enquiry in the concept of reasonableness), instead the maxim simply reflects a social virtue that enables us to understand our moral intuitions.

In this chapter, I have argued that if the concept of reasonableness is going to be the basis of the contractualist theory of reasonable agreement, it is necessary to search for a definition that divorces reasonableness from our moral motivations. The importance of such an account of reasonableness is that it provides an independent standard or rationale to illuminate our moral intuition to seek agreement.

The point is that moral motivations taken in their isolation are susceptible to two criticisms: they are either predominantly psychological attributes, in which case these motivations are devoid of a political dimension and therefore insufficient as a foundation of a moral theory, or they are philosophical accounts of moral psychology drawn from political conceptions, in which case moral motivations are so well formulated⁹⁰ to make constructivism in ethics totally redundant.

I believe that we should be somewhere in the middle of these two positions, that is to say, the psychological aspect of our moral motivations should not be undervalued, although at the same time these should not become a mere reflection of our political conceptions. Instead moral motivations should be considered within a social and political context. It is the combination of our moral motivations and the social context that transforms such motivations into the foundations of a political theory.

The only way to tread this middle road is by rejecting the

⁹⁰ - By this I mean that they are not simply our unconsidered intuitions or judgments, but judgments we hold after due reflection.

view that the concept of reasonableness refers to our moral psychology. Instead I have argued that reasonableness ought to be seen as a rationale for illuminating our moral motivations. I believe that the concept of reasonableness, defined in terms of the maxim 'not asking for too much', provides the grounds on which our moral motivations are to be tested.

PART II - REASONABLE AGREEMENT AND POLITICAL LIBERALISM

It is a well-known fact that the aftermath of any revolution is demarcated by two factors. First, the extensive fragmentation of political units fighting for legitimacy and supremacy. Secondly, the success of any revolution is determined by its ability to replace existing institutions with alternative ones which are both stable and legitimate.

While most people would not hesitate to accept these two rules as being relevant to history and politics, perhaps only a few realize that the same rules apply to philosophy. Most political philosophers would acknowledge that 1971 represents a revolution in political philosophy. As Brian Barry points out, Rawls's A Theory of Justice is the watershed that divides the past from the present, indeed in political philosophy there is a pre-Rawlsian world, and a post-Rawlsian world (Barry 1990c).

In the three chapters that make up Part I, I focused on the revolutionary nature of Rawls's theory, and the fragmentation which resulted from it. A revolution signals a radical break with the tradition legitimizing the status quo, and if we consider

utilitarianism the dominant tradition in the pre-Rawlsian world, then contractualism represents the commanding credo in the post-Rawlsian world.

The fragmentation which characterised the immediate aftermath of the Rawlsian revolution has not spared the idea of contractualism, in fact we have seen how since 1971 opposing contractualist traditions have developed. On one side we find the contractualism of Scanlon, built around the idea of reasonable agreement, while on the other hand we find the contractualism of Gauthier, built around the idea of rational bargaining⁹¹.

The aim of my thesis so far has been to show that reasonable agreement is where the future of the Rawlsian revolution lies. I have argued that to the extent that it advocates a superior conception of equality, only the contractualist theory of reasonable agreement represents a valid alternative to utilitarianism. Indeed the theory of reasonable agreement draws out more clearly than Rawls's original position the moral intuition of impartiality which is the driving force behind Rawls's theory of justice.

In the chapters that follow, I will concentrate on the second factor that applies to all revolutions, namely, that a revolution is successful if and only if the principles it champions constitute a coherent and workable alternative. Thus it is not sufficient to

⁹¹ - The fragmentation in the post-Rawlsian world is not limited to contractualism, in fact a plethora of non-utilitarian political theories has emerged. Amongst the most influential alternatives we find libertarianism, perfectionism, communitarianism and analytical marxism.

eradicate the fundamental principles behind the existing status quo and replace them with a new set of principles. It is also necessary to show that the new principles being advocated are capable of acting as the foundation of a comprehensive political theory.

In the case of contractualism, by a 'comprehensive political theory' I am referring to the key notions that constitute the subject of political liberalism: political obligation, social justice and neutrality. The challenge facing contractualism now is exactly that of inferring from the basic principles of reasonable agreement theories of political obligation, social justice and neutrality, and to combine these three key concepts in a comprehensive theory.

It is important to realize that undermining the inadequacy of the utilitarian account of one of the three key notions of political liberalism is not sufficient to dethrone utilitarianism. One of the great assets of utilitarianism is that it is not a theory of political obligation, neutrality or social justice, instead it is all three at the same time. Utilitarianism manipulates political liberalism through its theory of moral motivation.

On the basis of the argument presented in Part I that reasonable agreement gives an alternative account of moral motivation to utilitarianism, in the chapters that follow I will argue that from the contractualist theory of reasonable agreement we can extract theories of political obligation, social justice and neutrality.

5. POLITICAL OBLIGATION

Do citizens have an obligation to obey the laws of the state? And if so, why? On what grounds is our duty to obey the law justified? The above questions set the agenda for the problem of political obligation, and for the last 350 years liberalism has been concerned to give persuasive answers to such questions.

The various attempts to construe a liberal account of political obligation have hitherto received a mixed reception. Indeed only a few years ago it was fashionable to claim that liberalism could not provide a workable theory of political obligation: two of the best known advocates of this position are Alan J. Simmons (1979) and Carole Pateman (1985). While some liberal thinkers have responded to this critique by denying that the citizen has an obligation to obey the law, hence denying political obligation *tout court*⁹², others have been seeking a reconciliation between liberalism and the problem of political obligation.

Among the latter group, two recent theories deserve closer inspection. George Klosko (1992) attempts to ground political

⁹² - See for example Raz (1986) and Green (1990).

obligation on the principle of fairness, while Jeremy Waldron (1993) argues that a natural duty to justice is the basis for a liberal theory of political obligation. Although the theories of Klosko and Waldron stand out as the most intelligible attempts to construe a liberal theory of political obligation, in what follows I will argue that the theories of Klosko and Waldron are both inadequate.

The validity of a liberal theory of political obligation can be tested in terms of its ability to respond to two questions. First, how the theory deals with the distinction between presumptive and discretionary goods, that is to say, does the theory of political obligation justify a duty towards a state providing only presumptive goods, or does it also justify a duty towards a state which provides also discretionary goods? Secondly, how a theory reconciles our duty towards the state with our right to civil disobedience, that is to say, is a theory of political obligation capable of distinguishing between an overall obligation to a just state and a non-obligation towards unjust single laws⁹³? Throughout this chapter, I will refer to the combination of these two questions as the validity-test.

The argument in this chapter is developed in four main parts. Parts 5.I and 5.II will concentrate respectively on the theories of Klosko and Waldron. I will argue that both theories fail the

⁹³ - Richard Bellamy (1994) believes that the question of civil disobedience creates fundamental problems for contemporary liberal theories, in fact liberalism is incapable of reconciling theories of political obligation with accounts of civil disobedience.

validity-test, hence their attempts to construe a liberal theory of political obligation is not totally successful. Thus while Klosko's theory is incapable of giving adequate responses to either question, Waldron's theory has difficulty with the latter question.

In Part 5.III, I will put forward an alternative theory of political obligation which I believe passes the validity-test. I will argue that this theory finds justification in the contractualist theory of reasonable agreement.

Finally, I will consider the views of Bhikhu Parekh on political obligation. Parekh's theory is interesting because it appears to clear the obstacles set by the validity-test, furthermore he approaches the question of political obligation from a moral angle. In Part 5.IV, I will argue why we should resist the temptation to follow Parekh's lead on questions of political obligation.

5.I. Fairness and Political Obligation.

George Klosko claims that the materials for a solution to the problem of obligation have been at hand since 1955, when H.L.A. Hart's essay "Are There Any Natural Rights?" was first published. The principle of fairness was originally formulated by Hart as follows:

[W]hen a number of people conduct any joint enterprise

according to rules and thus restrict their liberty, those who have submitted to these restrictions when required have a right to a similar submission from those who have benefited by their submission. (Hart 1955, p.185; quoted in Klosko 1992, p.33)

This principle has since been adopted by Rawls, although not in the context of political obligation. Klosko's contribution was to apply Hart's principle of fairness to the problem of political obligation.

The moral intuition behind the principle of fairness is that it is wrong to 'free-ride' on the efforts of others. On a more basic level, the principle of fairness expresses the general idea that similar individuals should be treated similarly, since it is wrong for certain people to be exempt from burdens others must bear in the absence of morally relevant differences between them (Klosko 1992, p.34).

Klosko's attempt to ground political obligation on the concept of fairness has some advantages but also some devastating disadvantages. One advantage is that grounding political obligation on the principle of fairness helps to overcome the kind of problems that have inflicted consent-based accounts of political obligation. For example A.J. Simmons argues that consent theories are not able to account for the political obligation of large number of citizens, yet as Klosko points out: "The principle of fairness does not ground obligation on consent - either express or tacit - and so

its generality is not hampered by the need for citizens to consent to their governments" (Klosko 1987, p.356).

Another advantage of grounding political obligation on the principle of fairness is that it provides a moral standard from which to assess our sense of duty; thus political obligation is not based on the idea of mutual advantage or on a right to benefit from co-operation, but on a duty to reciprocate.

Notwithstanding the above advantages, there are two major disadvantages associated with the idea of grounding obligation on fairness, which threaten to undermine the validity of Klosko's theory. First, it fails to distinguish between presumptive and discretionary goods, and secondly, it fails to take into account cases of civil disobedience. I will consider these two issues separately.

Let us begin by defining the terms in question. Presumptive public goods are goods that are indispensable, that is to say, they are necessary to all members of the community regardless of their individual conception of the good. For example law and order are presumptive goods in a way that a national art gallery is not. Discretionary public goods are goods that are desirable rather than indispensable. For example, state support for the arts and the reparation of roads are desirable although strictly speaking not essential, hence these goods are discretionary⁹⁴.

The distinction between presumptive and discretionary goods

⁹⁴ - The distinction between presumptive and discretionary goods can be compared to Scanlon's (1975) distinction between urgency and preferences.

has important political implications. For example Nozick's minimalist state is a state that provides exclusively presumptive goods, while Rawls's principles of justice promote a society where the state provides both presumptive and discretionary goods. Unless one wants to embrace a libertarian philosophy, and I don't think this is Klosko's intention, it is necessary to show that we have an obligation towards a state that supports the provision of discretionary goods.

Yet it appears that Klosko's theory of political obligation, grounded on the principle of fairness, fails to justify a duty towards a state that provides discretionary goods. Indeed Klosko seems to be aware of this deficiency:

Though [the principle of fairness] creates strong obligations for individuals to contribute to the provision of presumptive goods, it does not create obligations to help provide discretionary public goods. (Klosko 1992, p.85)

The problem Klosko faces is that as it stands, the principle of fairness seems able to justify obligations only to support a minimalist state⁹⁵, not a state that supports the provision of discretionary goods. Not being sympathetic to a Nozickian society, Klosko is keen to convince us that this objection can be overcome, and he devotes a whole chapter of his book to a discussion of this

⁹⁵ - That is to say, a state providing only presumptive public goods.

problem.

Klosko's solution to the problem is the following: one needs to extend the functions of cooperative schemes that have been set up to provide presumptive public goods in order to include also discretionary goods. In other words although the principle of fairness cannot oblige individuals to support schemes that furnish only discretionary public goods, it can generate obligations to a scheme that provide discretionary goods if it can be shown that the given scheme also provides presumptive public goods (Klosko 1992, p.86).

I find Klosko's response unsatisfactory, indeed Klosko's failure to deal with the question of discretionary goods adds support to the claim that a theory of political obligation grounded on the principle of fairness is valid only to justify a minimalist state. In what follows I will first look at Klosko's solution in detail, followed by an account of why his response is far from being convincing.

Klosko argues that two main considerations support obligations to contribute to discretionary public goods if they are provided by cooperative schemes that also provide presumptive public goods: the indirect and the institutional argument. The general idea behind the indirect argument is that presumptive and discretionary goods comprise an indivisible benefit package; it is because discretionary public goods fall in such package that individuals can be obligated to support them. The institutional argument is similar to the indirect argument. It states that presumptive and

discretionary goods are part of the same decision process, hence of the same institutional package. It follows that even if institutional benefits (such as law and order) are closely related to presumptive goods, they extend to discretionary goods as well.

I believe that both the indirect and the institutional argument fail to bail out Klosko from the charge that grounding political obligation on the principle of fairness only justifies a duty to a minimalist state, in fact Klosko seems unable to justify our obligation to a state that provides discretionary goods.

I suggest we start from the indirect argument. The point of the indirect argument is that the provision of discretionary public goods is instrumental for the provision of presumptive public goods. As Klosko points out:

The indirect argument relies on the *practical* indispensability of certain discretionary public goods In other words, though discretionary public goods are not directly required for individual well-being, they are required indirectly.

(Klosko 1992, p.87)

Klosko's example of an indivisible benefit package is that of a presumptive good (national defense) and a discretionary good (repairing roads); according to Klosko, some level of transportation and other discretionary public goods are indispensable to national defense - and to the provision of other presumptive public goods. I think Klosko's indirect argument fails

on two accounts.

First of all, the example of national defense and repairing roads is misleading. If Klosko wants to argue that the provision of discretionary public goods is instrumental for the provision of presumptive public goods, I suggest he explains how the provision of discretionary goods as national art galleries are indispensable to national defense. Such argument would need to show that if all soldiers were erudite in 16th century Flemish paintings, they would be better soldiers. Although this makes for a fascinating experiment, I am sceptical that it can produce a valid result, therefore we can safely conclude that the idea of an indivisible benefit package is unpersuasive.

Secondly, and more seriously, it seems to me that Klosko has not responded to the challenge of showing how the principle of fairness can provide an obligation for discretionary goods, instead all he has done is to dissolve the distinction between presumptive and discretionary goods by defining away the problem. Klosko tells us that an objection along these lines was brought to his attention by B.J. Diggs, and Klosko devotes a one-page-long endnote trying to respond to the charge that by making discretionary goods indispensable the distinction between presumptive and discretionary goods is obfuscated to the degree of being meaningless. Klosko's reply is that presumptive and discretionary goods differ in two main respects: (1) the kinds of uses to which they are put; and (2) the kinds of problems they are designed to overcome. I believe both arguments can be refuted.

Klosko argues that presumptive goods have more specialized uses, while discretionary goods have more varied uses. For example while the components of the criminal justice system - police, jails, courts, parole boards, etc. - are closely tied to the supply of law and order (a presumptive good), the provision of transportation and communication facilities (discretionary goods) are "components of an overall societal infrastructure and so less closely bound up with specific indispensable functions" (Klosko 1992, p.109n13).

I cannot see how appealing to degrees of specialization helps Klosko to confront the accusation of defining away the distinction between presumptive and discretionary goods. The question here is not the degree to which these types of goods are indispensable, but whether they are indispensable or not. By maintaining that presumptive and discretionary goods are part of the same "overall societal infrastructure" Klosko is hammering the last nail in his coffin, since the idea of an overall societal infrastructure totally undermines any hope of distinguishing between different types of public goods. By being part of the same overall societal infrastructure, discretionary goods and presumptive goods are both indispensable, even if in different degrees, hence their distinction disappears.

The second major difference between presumptive and discretionary goods is that they address different kinds of problems: "In sum, presumptive public goods alleviate specific, pressing problems, while the difficulties that discretionary public

goods confront are more diffuse and less immediate, if no less real" (Klosko 1992, p.110). Klosko maintains that the difference between the two types of goods boils down to a difference in time frame, thus while failure to provide presumptive goods leads to an immediate catastrophe, the result of not providing discretionary goods would be no less catastrophic, but a longer time frame would be involved.

Once again, it seems to me that Klosko fails to respond to the accusation that he obfuscates the distinction between discretionary and presumptive goods. The fact is that in Klosko's account the provision of presumptive and discretionary goods are similar to the extent that they both help to prevent catastrophes, and the issue of a longer or shorter time frame does not hide the fact that they are both similarly indispensable.

As Klosko's many arguments fail, one is left with the impression that by grounding political obligation on the principle of fairness Klosko is incapable of justifying a duty towards a non-minimalist state, that is to say, a state that provides both discretionary goods and presumptive goods. Yet this is not the only deficiency in Klosko's theory, in fact Klosko is also incapable of reconciling a theory of political obligation with the right to civil disobedience. It is to this question that I want to turn to now.

So far I have rejected Klosko's claim that the *indirect* argument explains how the principle of fairness justifies our obligation towards a state providing discretionary public goods. Of

course this is not Klosko's only argument, in fact (presumably in case the reader is not convinced) Klosko puts forward an alternative argument which he thinks is capable of redeeming his theory from the accusation of justifying only a duty towards a minimalist state, namely, the *institutional argument*.

The thrust of Klosko's institutional argument for political obligation is that disobedience has potentially devastating effects:

The main reason why individuals must obey the law is the corrosive effects of disobedience. If A disobeys the law, he acts to undermine the rule of law, which is central to the welfare of society and to his own welfare as well. (Klosko 1992, p.101)

In other words there is no difference between an obligation to follows rules that provide presumptive goods (i.e. law and order), and an obligation to follow rules that provide discretionary goods (i.e. road repairs), since all rules are based on the same institution, namely, the rule of law: "Thus 'A' should not disobey minor laws; because general disobedience of even minor laws would undermine law and order throughout society, it would be unfair of him to disobey them" (Klosko 1992, p.103).

I believe Klosko's institutional argument is seriously defective. The most unattractive aspect of Klosko's institutional argument is that it rejects all claims for civil disobedience,

indeed Klosko's institutional argument is incapable to even acknowledging cases of civil disobedience.

It is important to remember that civil disobedients, as Richard Bellamy rightly points out, are neither revolutionaries aiming at total social disintegration, nor criminals concerned only with personal gain, instead civil disobedients "accept the legitimacy of the existing legal order and wish merely to amend a particular law" (Bellamy 1994, p.26). Civil disobedients, Bellamy continues, "frequently claim that because the particular law they are protesting about is at odds with the norms and practices of the polity as a whole, it is illegitimate and likely to bring the whole regime unjustifiably into disrepute" (Bellamy 1994, p.26).

Judging from Klosko's definition of the institutional argument, it would appear that according to Klosko an act of civil disobedience is as criminal as an act of terrorism. If, as Klosko points out, "the institutional argument supports obligations to obey all laws" (Klosko 1992, p.105)⁹⁶, this includes an obligation to follow laws that may be regarded unjust.

Klosko attempts to counter this further complication by considering a counterexample, although as we shall see this is unsatisfactory for a number of reasons. The example Klosko considers concerns adhering to the laws of a state "as grievously unjust as Nazi Germany" (Klosko 1992, p.103). In this case Klosko argues that we don't have an obligation to obey such laws since

⁹⁶ - As Klosko explains: "Because of the possibility that violating any law will erode one's habit of obedience, there is a *prima facie* obligation to obey all laws." (Klosko 1992, p.105).

this obligation would be overridden by the moral requirement not to pursue unjust acts. There are two major problems with Klosko's example. First, Klosko is appealing to a further moral principle (for example that it is wrong to commit unjust actions or support their perpetration), which has little in common with the principle of fairness. It seems to me that appealing to a further moral principle undermines the utility and validity of the principle of fairness; I shall return to this point later.

The second reason why Klosko's example is unsatisfactory is because it fails to address those hard cases that would really stretch his theory of political obligation. Consider the example of Great Britain in the 1980's and early 1990's, a state that is sufficiently just to be considered legitimate. Apart from providing presumptive goods, the government of Great Britain passes or preserves laws that many feel to be unjust: examples are the poll tax, the T.V. licence, the Road Tax, or adding 17.5% VAT on fuel. All four examples have in common the fact of being flat taxes.

In those cases, a citizen may feel that he or she has a general obligation to the state, but not towards the single issues he or she regards grossly unjust. Thus for example I may feel I have an obligation to pay my income tax and respect criminal laws, and perhaps even join the army if the country needs me, nevertheless I may also feel that the poll tax is unjust. Am I under an obligation to pay it? According to Klosko, unquestionably 'yes'.

The problem I have with such unconditional response is that it

leaves no room for civil disobedience. This is a clear indication that Klosko's theory is clearly unable to deal with the second question in the validity-test, namely, being able to distinguish between our obligation to obey the laws of a just state, and a non-obligation to obey single laws deemed to be unjust.

Considering the considerable difficulties Klosko has in responding to the validity-test, it is perhaps not surprising that on a number of occasions (as in the example of Nazi Germany) Klosko admits that fairness theory should be supplemented with other moral principles, especially a duty to justice (Klosko 1992, p.91 and p.98)⁹⁷. What Klosko fails to realize is that this concession threatens to undermine the validity of his theory of political obligation.

Surely a moral principle that can justify the provision of discretionary goods as well as allow for cases of civil disobedience, can also justify the provision of presumptive goods and a general respect for the laws. It follows that this other moral principle, whatever it may be, is the real foundation behind

⁹⁷ - Without referring to Klosko's work, Richard Bellamy reaches the same conclusion, namely, that fairness theory requires a substantial account of justice to ground it. According to Bellamy, it is essential to the fair-play thesis that the scheme of co-operation be fair and the benefits be accepted, yet it is highly debatable that agreement can be reached on what counts as fair: "Rival accounts can be given, for example, from the perspective of desert, need and entitlement, and each of these alternatives are open in their turn to differing interpretations ... [W]ithout a substantial account of justice to ground it, the fair-play argument will be unable to provide an adequate account of political obligation" (Bellamy 1994, p.38).

political obligation⁹⁸, indeed it makes the principle of fairness entirely redundant.

If fairness theory applies only to simple cases, but requires supplementation by other moral principles when confronted with hard cases, then the principle of fairness cannot be the foundation for a liberal theory of political obligation. Instead we would be better advised to look at the other moral principles that supplement the principle of fairness for the true foundation of political obligation.

So far I have argued that an approach to political obligation based on the principle of fairness fails the validity-test. In Part 5.II, I will argue that Waldron's theory of a natural duty to justice provides a better theoretical foundation of political obligation than the principle of fairness, although it is not as convincing as the solution I favour, namely, reasonable agreement.

5.II. Natural Duty to Justice and Obligation.

In "Special Ties and Natural Duties" Waldron (1993) argues that there are two accounts of political obligation: theories of acquired obligation and theories of natural duties. So far liberal political theorists have neglected theories of natural duties while focusing almost entirely on theories of acquired obligation, and as

⁹⁸ - In Part 5.III, I will argue that the notion of reasonable agreement provides the basis for such moral principle.

a result the debate has been between two approaches to acquired obligation: theories of consent and fair play.

Waldron wants to resurrect the theory of natural duty to its rightful place at the centre of the debate on political obligation. In what follows, I will first give a brief summary of Waldron's theory of political obligation as natural duty to justice, followed by an account of how his theory fares in terms of the two questions that make up the validity-test.

According to Waldron, the claim that we have a natural duty to support the laws and institutions of a just state is less well known, indeed he tells us that there is less written on it. Waldron finds this surprising since he claims that philosophers toy with something like the theory of natural duty in almost all their theories about what people owe to the state.

It may be argued *contra* Waldron that his attempt to dichotomise the accounts of political obligation into two separate camps is illegitimate, since all theories based on consent or fair play must necessarily assume natural duties, and vice-versa⁹⁹. For example Rawls's original position is an obvious attempt to device the conditions for an hypothetical consent, indeed in a different article Waldron rightly points out that according to Rawls: "The test of a just society, then, is not whether the individuals who live in it have agreed to its terms, but whether its terms can be represented as the object of an agreement between them." (Waldron

⁹⁹ - Unless by consent theories Waldron is thinking of only actual consent theories, in which case he is challenging a straw man.

1987, p.142, emphasis in original).

The point is that (hypothetical) consent theories are more interested in establishing the terms of a just society, rather than portraying the feelings of the individual subjects. Similarly theories of natural duties are also interested on the terms of a just society: saying that citizens have a duty to support the laws and institutions of a just society, or that citizens would consent to support the laws and institutions of a just society, amounts to the same thing. Hence it appears that there has always been an overlap between theories based on consent or fair play and theories based on a natural duty to justice.

Yet my major concern with Waldron's approach has little to do with his labels, in fact even if we accept Waldron's categories, and grant theories of obligation based on natural duties autonomous status, I feel there are still some major problems with his thesis; in what follows I will argue that Waldron's approach fails because of its lack of a meta-ethical dimension, which creates special problems especially when Waldron's theory of political obligation comes to face issues of civil disobedience.

The backbone of the natural duty position is that our cooperation in establishing and sustaining political institutions that promote justice is morally required. In order to sustain this thesis, Waldron points out that justice is partly a matter of cooperation, and partly a matter of coordination. The pursuit of justice requires that people cooperate together, and for the network of cooperation to be coordinated by institutions.

Since the avoidance of injustice is a moral imperative, it follows that the establishment of coordinating institutions is a moral imperative, which means that an institution that fosters cooperation and coordination can impose itself on us on moral grounds, indeed justifying such imposition is the strength of a natural duty approach: "The assumption of the natural duty approach is that the pursuit of justice is a moral imperative" (Waldron 1993, p.28).

Armed with the above account of Waldron's theory of political obligation, we can now evaluate how this theory fares in terms of the validity-test. It seems to me that Waldron's theory is to be preferred to Klosko's on at least one points, in fact Waldron's natural duty to justice approach is capable of justifying a duty towards a state that provides discretionary goods. Yet Waldron's theory is unduly uncompromising in terms of cases of civil disobedience, hence it fails the second trial set by the validity-test.

By appealing to our natural duty to support the laws and institutions of a just state, Waldron may be able to justify a duty towards a state that provides discretionary goods. In fact all Waldron has to do is to argue that the provision of discretionary goods is part of what is considered a just society. Waldron's idea of a just society is that of a co-operative venture, hence the provision of discretionary goods is needed in order to foster the spirit of cooperation.

While it cannot be denied that a society is a co-operative

venture, I believe there is more to a just society than co-operation. In other words, it is not co-operation *per se* that is important for a just society, instead what matters are the *terms* of cooperation¹⁰⁰. I believe fair terms of cooperation can be established only by a meta-ethical theory, which is something totally lacking in Waldron's theory.

It follows that the problem with Waldron's theory of political obligation is that it justifies an obligation to a state *tout court*, that is to say, it fails to take account of cases where people want to express a dissatisfaction with a specific law without necessarily undermining the legitimacy of the state. To recapitulate the argument, I believe there are two problems with Waldron's approach. First, by defining a just society in terms of a co-operative venture, Waldron has no choice but to reject non-cooperative actions as unjust. This would include the anti-poll tax non-payment campaign, where non-payers would be accused of free riding on the co-operative efforts of others.

Secondly, Waldron's approach to the problem of political obligation relies heavily on his theory of political legitimacy:

The natural duties come into play only where the organization in question passes not only tests of justice and effectiveness, but also a test of legitimacy. (Waldron 1993, p.22)

¹⁰⁰ – Indeed according to Rawls, fair terms of cooperation are the foundations of a well-ordered society.

yet as we shall see his theory of political legitimacy is not equipped to deal with specific issues of civil disobedience.

Waldron tells us that political legitimacy has to do with "strategic choice in something like a coordination game" (Waldron 1993, p.25). It seems to me that this approach to political legitimacy has a strong functional flavour, which risks to overshadow the important distinction between just and unjust laws within a just society. In other words, a system of rules may be legitimate, while at the same time a particular law may be illegitimate.

To conclude, I am not denying that it is possible to ground political obligation on the idea of a natural duty to justice. I just think that this theory of political obligation is not sufficiently thorough. In particular I am sceptical that this approach can ever reconcile political obligation with civil disobedience. The natural duty approach is an 'everything-or-nothing approach'; unless Waldron supports his theory of a natural duty to uphold justice with a meta-ethical account of the just society, his theory of political obligation will always lack the dimension necessary to deal with issues of civil disobedience.

5.III. Obligation and Reasonable Agreement.

So far I argued that Klosko's approach, based on the principle of fairness, fails the validity-test because of its inability to

justify a duty towards a state providing discretionary goods, as well as being unable to reconcile our obligation towards a just state, with our right to oppose unjust laws. I have also argued that Waldron's approach, based on the idea of a natural duty to justice, fails to deal adequately with questions of civil disobedience.

I believe that a theory of political obligation based on reasonable agreement, where reasonableness is defined by the proviso 'not asking for too much', can pass the validity-test by providing adequate responses to both questions. Before addressing the problem of political obligation, it will be necessary to provide a brief restatement of the notion of reasonable agreement.

In Chapter 3, I argued that in the contractualist theory of reasonable agreement our moral motivations are defined by the idea of agreement, not the concept of reasonableness. The idea of agreement states that in seeking an agreement, we are willing to compromise the pursuit of our conception of the good in order to make room for others¹⁰¹. This idea of agreement has a strong affinity with the principle of impartiality, where impartiality is defined in terms of the belief that the well-being of all moral beings matters intrinsically¹⁰². In other words the idea of agreement substantiates the recognition of others as having

¹⁰¹ - M. Gilbert (1993) argues that although the parties of an agreement are viewed as separate individuals, by entering an agreement we thereby constitute ourselves the members of a (collective) *we*; an agreement is the foundation of mutual recognition and joint commitment.

¹⁰² - See Kymlicka (1990b), p.111.

legitimate claims to have their interests taken into account¹⁰³.

Another key aspect of the theory of reasonable agreement is that the concept of reasonableness is not the key to our moral motivations, instead the concept of reasonableness acts as the rationale that illuminates and helps us to understand our moral intuitions. Thus in Chapter 4, I argued that reasonableness ought to be defined in terms of the maxim 'not asking for too much'.

Starting from this definition of reasonableness, we can define the notion of reasonable agreement as an agreement based on the desire to seek an agreement, where such desire is qualified by the requirement that we should not be 'asking for too much' from others.

Now that we have a general idea of the notion of reasonable agreement, we must inquire into the ways in which this notion relates to the question of political obligation. I think we find the answer to this question, albeit at its embryonic stage, in the following claim by Scanlon:

An act is wrong if its performance under the circumstances would be disallowed by any system of rules for the general regulation of behaviour which no one could reasonably reject as a basis for an informed, unforced general agreement.

(Scanlon 1982, p.110)

According to Scanlon, there is a correlation between an act

¹⁰³ - See Barry (1989b), p.269.

that is morally right, and a system of rules which no one could reasonably reject. What this means, in terms of political obligation, is that we have a duty towards a system of rules that is the outcome of a reasonable agreement, or an agreement that no one could reasonably reject.

For a theory of political obligation based on reasonable agreement, it is important that the system of rules reflects a commitment to impartiality, whereby the equal consideration of interests of all is defended. The duty of the citizen is directly proportional to the impartiality of the system of rules; if impartiality is not enforced by the state institutions, the conditions for an hypothetical agreement are violated, and as a result political obligation is undermined.

Our political obligation is directly related to the neutrality of the state, where neutrality means that the system of rules as a whole is based on a set of values that no one could reasonably reject as the basis of an agreement. It is important to emphasize that while we have an obligation to the system of rules as a whole, we do not necessarily have an obligation to single rules¹⁰⁴.

The reason why I feel that Scanlon's claim is only a partial answer to the problem of political obligation, is because Scanlon fails to specify what he considers to be a 'reasonable rejection', indeed we have seen that Scanlon eschews all attempts to define the

¹⁰⁴ – For a detailed account of neutrality, see Chapter 7 below: 'On Liberal Neutrality'.

elusive concept of reasonableness¹⁰⁵.

By defining the concept of reasonableness in terms of the maxim 'not asking for too much', I believe we are in a position to deduce from the notion of reasonable agreement a more precise account of political obligation. Briefly, the argument runs as follows: we have an obligation towards the state if the state endorses a system of rules that is not asking for too much from its subjects. The idea of agreement, combined with the maxim 'not asking for too much', reflects a concern for the consequences of our actions on others. This concern is the core intuition behind the principle of impartiality, since being concerned with the effect of our actions on others enforces the principle of equal consideration for other people's interests.

In what follows, I will argue that a theory of political obligation based on the notion of reasonable agreement is capable of satisfying the requirements set by the validity-test, in fact it justifies our duty towards a state that supports the provision of discretionary goods, and it reconciles political obligation with the principle of civil disobedience.

In order to demonstrate that a theory of political obligation grounded on the idea of reasonable agreement is able to justify a duty towards a state that provides for discretionary goods, the first step will be to show that both discretionary and substantive goods are instrumental towards forming a political community based on a system of rules that no one could reasonably reject. In other

¹⁰⁵ - See Chapter 4, Part I.

words, a state that provides exclusively substantive goods could be reasonably rejected by a faction of citizens.

The key to the idea of reasonable rejection lies in the principle of impartiality. Thus a state that provides only substantive goods inevitably violates the principle of impartiality. If impartiality implies that the well-being of all moral beings matters intrinsically, hence that everyone has a legitimate claim to have their interests taken into account, then by not providing for discretionary goods the state is in fact transgressing the interests of those members of a polity that are incapable of pursuing their conception of the good.

Another way of expressing this concept is by saying that unless a state provides for discretionary goods, some members of the polity will be disadvantaged by morally arbitrary circumstances, such as the natural and social lottery. Furthermore for the theory of reasonable agreement, those who would stand to gain from a state that provides only substantive goods are not in a position to reasonably reject a state that provides both substantive and discretionary goods, even if this would mean that they are worse off.

Consider the following example. State A provides only substantive goods, while State B provides both discretionary goods and substantive goods. As a citizen of State A, Margaret (who is relatively well-off) would enjoy a living standard of X, while under State B Margaret would enjoy a living standard of Y, where $Y < X$. The difference between living standards X and Y is the amount

Margaret has to contribute to State B in order for it to provide discretionary goods. In other words, $X - Y =$ the difference in level of taxation between State A and State B.

According to the theory of reasonable agreement, Margaret has a political obligation towards State B, or in other words, Margaret cannot reasonably reject the system of laws of State B. Rejecting State B on the grounds that one is better off under State A contradicts the very idea of agreement, since according to this moral motivation Margaret must be willing to compromise the pursuit of her conception of the good in order to make room for others.

Although Margaret is better off under State A than State B, she cannot claim that state B is asking for too much from her, since the maxim 'not asking for too much' only applies to people who seek an agreement by endorsing the principle of impartiality, which is not the case of Margaret. The inverse is in fact closer to the truth. That is to say, all those that would benefit from State B can argue that under State A people like Margaret are in fact 'asking for too much' from them, since they cannot pursue their conception of the good (or only under considerable sacrifice) while people like Margaret can do so with ease.

So far I have argued that a theory of political obligation based on reasonable agreement would justify an obligation towards a state providing both substantive and discretionary goods, while not necessarily justifying an obligation towards a minimalist state providing exclusively substantive goods. I want to consider now the case of civil disobedience.

I believe it can be shown that a theory of political obligation grounded on the notion of reasonable agreement does not undermine the legitimacy of civil disobedience. That is because reasonable agreement is capable of distinguishing between an obligation to a just state, and an obligation to a particular law, which may be unjust, within a just state.

By grounding political obligation on the theory of reasonable agreement we can allow for civil disobedience in cases such as the anti-poll tax campaign. The reason why the poll tax was susceptible to civil disobedience is because it was not proportionate to income. It is this feature of the tax that comes into conflict with the notion of reasonable agreement; by unfairly penalizing some individuals more than others, a flat tax contradicts the foundations of reasonable agreement¹⁰⁶.

The point to stress here is that one could reasonably reject to the poll tax on the grounds that it was based on a principle which is irreconcilable with impartiality. By imposing a flat tax the British state was in fact discriminating against the less well off, making it even more difficult for them to pursue their conception of the good.

Although in theory the benefits of paying the poll tax are shared by everyone in the community¹⁰⁷, a flat tax is asking for

¹⁰⁶ - I think the same conclusion applies to all forms of flat tax, for example the T.V. licence and road tax. Had the poll tax been a progressive tax, I am convinced that it would not have instigated the wrath of the British people.

¹⁰⁷ - Even if in practice increases in poll-tax were not proportionate to increases in local expenditure.

'too much' from those who are less wealthy compared to those who are more wealthy for the use of the same services. As Bellamy points out, the poll tax was held to be related directly to the benefits obtained from the services consumed, yet "research in this area shows that the poll tax actually had the effect of substantially increasing the relative benefit shares of upper income groups and lowering that of the poorer groups" (Bellamy 1994, p.40). It is important to remember that households with an average weekly income of between £50 and £200 were on average net losers under the poll tax, while the rest were on average net gainers. Furthermore while those with an average weekly income of under £50 gained on average by 0.01%, those with an average weekly income of over £500 gained on average by over 6.73 %¹⁰⁸.

Grounding political obligation on the notion of a reasonable agreement, where reasonableness is defined as 'not asking for too much', enables us to develop a double-standard, whereby one can express an obligation towards the just state (for example by paying income tax or peacefully accepting the results of a general election), while at the same time rejecting the obligation to follow single laws deemed unjust (the poll tax).

5.IV. Moral and Political Obligation.

In this chapter I have argued that the notion of reasonable

¹⁰⁸ - See Gibson (1990), Ch.5; quoted in Bellamy (1994), p.33n.

agreement provides persuasive answers to the two questions that make up the validity-test for theories of political obligation. Compared to the theories of Klosko and Waldron, grounding political obligation on the notion of reasonable agreement has two major advantages. First, it justifies a duty towards a state that provides both substantive and discretionary goods. Secondly, it reconciles political obligation with cases of civil disobedience, thus for example according to reasonable agreement it is not inconsistent for British residents to feel an obligation towards the British state, while at the same time feeling no obligation to pay the poll tax.

It must be said that grounding a theory of political obligation on reasonable agreement is not the only way to overcome the question of civil disobedience. Among contemporary political philosophers, Bhikhu Parekh (1993) is one of the few to have considered civil disobedience as a special case for political obligation, indeed he goes as far as to argue that civil disobedience is the very essence of political obligation. For this reason Parekh's theory deserves detailed consideration. Yet I believe that a closer inspection will show that Parekh's theory of political obligation is not as persuasive as it appeared at first, in fact Parekh redefines the concept of political obligation in such a way that its problematic nature is obfuscated.

At the basis of Parekh's argument we find two claims: first, that political obligation is not about obeying the civil authority or the law (he refers to these respectively as civil and legal

obligation), instead political obligation is about one's relationship to fellow-citizens collectively, or more specifically one's membership of a polity or community, the conduct of its collective affairs, and the preservation and enrichment of the quality of its collective life (Parekh 1993, p.243). Secondly, that citizens are above all moral agents, therefore capable of autonomous choice and responsible for the consequences of their actions (Parekh 1993, p.240).

On the basis of these two claims, Parekh builds a theory of political obligation which does not stop at a duty to obey the law. According to Parekh's theory, political obligation is about defending the interests of the moral community, even if this means breaking the law:

If they [citizens] feel convinced after calm reflection that a law confronts them with unacceptable demands or is likely to cause serious harm to the interests of the community, they have an obligation to criticize and protest against it and even perhaps disobey it. Such disobedience temporarily suspends their legal obligation in a specific area and affirms their status as moral agents. (Parekh 1993, p.240)

The above claim shows that according to Parekh civil disobedience is at the very heart of political obligation in general, and of an obligation to support a moral community in particular:

I have tried to show so far, somewhat tentatively, that citizens have several obligations in addition to obeying the law. These include an obligation to take an active interest and to participate in the conduct of public affairs, to keep a critical eye on the activities of the government, to speak up against injustices of their society, to stand up for those too demoralized, confused and powerless to fight for themselves, and in general to help create a rich and lively community. (Parekh 1993, p.243)

There is no doubt that Parekh's account of political obligation is appealing. First of all, it takes seriously the question of political disobedience. Secondly, because Parekh builds his theory of political obligation around the question of civil disobedience, he can easily justify a duty towards a state that provides discretionary goods. The argument would run along the following lines: unless the state provides for discretionary goods, citizens have a duty to protest. In fact the provision of discretionary goods is essential in order to redress 'the injustice of society', to help those 'too demoralized, confused and powerless', and in order to create 'a rich and lively community'.

Notwithstanding the appeal of Parekh's theory, I feel there is a major weakness in his account, namely, it is not an account of political obligation, instead it is an account of moral obligation. Parekh distinguishes between legal, civil and political obligation, claiming that obeying the law is a legal obligation, while politics

and political obligation is about the community of moral agents. I believe this distinction is suspect.

The accounts of political obligation discussed in this chapter, as well as the majority of other theories of political obligation that I have failed to mention, are concerned with our duty to obey the law, yet Parekh relegates this to the secondary question of legal obligation. It follows that Parekh's account of political obligation has in fact very little to do with the rest of the literature on the subject.

Contrary to the views of Parekh, I feel that (a) we cannot make a distinction between legal, civil and political, since the overlap between these three spheres is too great; and (b) Parekh fails to specify with sufficient accuracy the relationship between political obligation and moral obligation. I will consider these two issues separately.

Concerning the first issue, Parekh fails to show that there is a difference between a legal and a civil obligation, and between the latter and political obligation. Thus Parekh claims that although legal obligation presupposes and is derived from civil obligation, these two types of obligations are not the same thing.

The example Parekh appeals to in order to make this point is that of wars or other types of national emergency, where the subjects are expected to give their full support and loyalty to the civil authority: Parekh concludes that "this has nothing to do with obeying the law". It seems to me that the example chosen by Parekh undermines the very distinction he is trying to make. When in the

1960's the young American boxer Cassius Clay (before he became Muhammed Ali) refused to be drafted to Vietnam, he was sentenced to jail for five months: it is because he refused to obey the law that he went to jail, not because he was expected to give full support to his civil authority. This example shows that Parekh has not been successful in distinguishing legal from civil obligations.

Similarly I believe Parekh fails to provide solid foundations for the subsequent distinction between civil and political obligation. He claims that citizens are primarily related to each other, and only secondarily to the civil authority (Parekh 1993, p.244), therefore citizens (political obligations) come before civil authority (civil obligations). Again I find this distinction hard to accept; to the extent that it is difficult (if not impossible) to distinguish the concept of citizenship from the context of civil authority in which it is defined, I cannot see how Parekh can claim that the obligation between citizens has priority over the obligation of the citizen to the civil authority.

Turning to the second issue, I find Parekh's treatment of political and moral obligation vague and ultimately unpersuasive. While the question of moral obligation is very important and perhaps slightly neglected by the literature, it is a different question from political obligation, dealing with different issues and focusing on different social agents. Theories of moral obligation are interested in the (horizontal) obligation of citizens to their fellow citizens, hence they are concerned with what contractualist authors referred to as the *pactum unionis*. On

the other hand theories of political obligation are interested in the (vertical) obligation of citizens to their rulers, hence they are concerned with what contractualist authors referred to as the *pactum subjectionis*.

Saying that moral obligation is not identical to political obligation does not mean that moral and political obligation cannot be reconciled, but it does mean that the relationship between politics and morality concerning questions of obligation must be the subject of a meticulous and careful examination. Unfortunately Parekh fails to provide this.

Parekh's main intention in his article is to distance political obligation away from legal obligation, and bring it under the wing of moral obligation. According to Parekh there is no reason why all obligations should be imposed by law, indeed Parekh claims that political obligation:

is a species of moral not legal obligation and falls outside the ambit of law. Political obligation is moral in its nature and form, political in its origin and content. (Parekh 1993, p.250)

Parekh's attempt to bridge the gap between the spheres of political obligation and moral obligation is interesting, yet I fail to see what he gains by rejecting any affinity with legal obligation. If Parekh's political obligation is not about legal obligation, then it cannot be compared to all the other theories of

political obligation, since in the last analysis Parekh is not confronting the question whether citizens have an obligation to obey the laws of the state. In other words by redefining the question of political obligation Parekh is avoiding all the major problems that inflict other theories of political obligation.

Furthermore, although Parekh wants to harmonize political obligation with moral obligation, he avoids tackling the key question: What is the relationship between morality and politics? Parekh's claim that "Political obligation is moral in its nature and form, political in its origin and content" is an ingenious way of dodging the issue. If political and moral obligation are to be reconciled, as I think they should, the only way of doing so is to grab the bull by the horns and work out what the moral nature of political obligation is, and how it relates to the political content.

In this chapter I have argued that in the contractualist theory of reasonable agreement, the idea of agreement and the maxim 'not asking for too much' provide the basis for a moral discourse, hence reasonable agreement is a sounder ethical foundation for a theory of political obligation. In order to support this claim I have argued that a theory of political obligation based on reasonable agreement does a better job of distinguishing discretionary from presumptive goods, and of explaining why (and when) we have an obligation to a non-minimal state.

6. SOCIAL JUSTICE: A COMPATIBILIST APPROACH

If there is a conclusive statement to be drawn from the many debates featuring social justice since 1971, it is that all theories of justice can be reduced to *a priori* moral intuitions. In fact it would be surprising if that was not the case. The very purpose of a theory of justice is to capture our moral intuitions about the concept of justice as the prime social value, and to formulate principles that determine the justice or injustice of situations and institutions¹⁰⁹.

The major problem political philosophers face, when dealing with the concept of social justice, is that it is impossible to make a direct equation between justice and our moral intuitions, since we are moved by a plethora of moral intuitions, and depending on the circumstances we appeal to one intuition rather than another. Indeed, it is not uncommon that we simultaneously appeal to more than one intuition. That is why any theory of justice that stems from any one intuition will always appear unsatisfactorily

¹⁰⁹ - For a useful discussion of the concept/conception distinction, see Campbell (1988), Chapter 1.

narrow and ultimately indefensible¹¹⁰. It follows therefore that, rather than argue for the superiority of one moral intuition over another, a theory of justice should ideally try to reconcile different moral intuitions.

In particular, I believe that two moral intuitions are at a basis of social justice: that it is unfair to suffer from ill-luck, and that it is unfair to be taken advantage of (i.e. have others free ride on your efforts). A theory of justice must be able to find room for both intuitions, and to resolve conflicts which arise when these two intuitions inevitably clash. Indeed the soundness of a theory of justice is directly proportional to its overall coherence after the reconciliation between compensating for ill-luck and rewarding merit has taken place.

These two fundamental moral intuitions are represented by two separate principles: respectively the principle of compensation and the principle of responsibility; these two principles will be analyzed in Part 6.I. In Part 6.II and 6.III, I will attempt to argue for a theory of justice that is conceived from the successful (unforced) marriage between the principles of compensation and responsibility. I will refer to this as the compatibilist theory of justice. In 6.IV and 6.V, I will argue that the compatibilist

¹¹⁰ - I believe this is the essence of James Fishkin's critique of systematic justice. A systematic theory of justice provides us (among other things) with an ideal which requires one or more first principles that hold without exception. This is the case of Rawls's theory of justice as fairness or utilitarianism. Fishkin argues that there are fundamental moral conflicts among competing first or ultimate principles, or in other words we have competing 'ideals without an ideal', hence no systematic theory of justice will ever be satisfactory. See Fishkin (1992), Part One.

theory of justice is not adverse to Rawls's theory of justice, indeed I believe that Rawls's second principles of justice exemplifies aspects of compensation as well as responsibility. Finally, in 6.VI I will argue that a theory of justice that reconciles the principles of responsibility and compensation would be endorsed by the contractualist theory of reasonable agreement.

6.I. What is Justice?

At least since Aristotle, it has become accepted that questions of social justice strike at the heart of egalitarian considerations¹¹¹. Echoing Aristotle's famous claim from the Politics that "all hold that justice is some kind of equality", Perelman claims that:

To everyone the idea of justice inevitably suggests the notion of a certain equality. From Plato and Aristotle, through St. Thomas Aquinas, down to the jurists, moralists and philosophers of our own day runs a thread of universal agreement on this point. The notion of justice consists in a certain application of the notion of equality. (Perelman 1963, p.12)

¹¹¹ - For a brief account of the correlation between justice and equality in the history of political philosophy, see Del Vecchio (1952).

The question for social justice is how to interpret and define the application of the notion of equality. Although to be just is to treat in equal fashion, it would be a mistake to assume that social justice can be reduced to the notion of treating everyone equally. As Dworkin and Kymlicka rightly point out, there are two ways of interpreting an egalitarian theory: in terms of treating everyone in the same way (for example by providing a perfectly equal distribution of income), or in terms of treating people as equals (for example acknowledging that each citizen is entitled to equal concern and respect). Theories of social justice are egalitarian in the latter sense.

Thus contemporary theories of social justice are concerned to explain and justify both equalities and inequalities; on the 'equality' side of this equation, all theories of justice must explain the concept of treating people as equals, while on the 'inequality' side of the equation, all theories of justice must justify legitimate grounds for differential treatment. The important point to emphasize is that all inequalities must be inferred from the initial postulate of equality. Thus for example Rawls starts from an account of an initial condition of equality (namely, the original position), and from this basic premise of equality he deduces principles that justify unequal treatment (the difference principle).

Starting from the assumption that social justice is concerned with the distribution of benefits and burdens, in what follows I will address two ways in which the egalitarian core of social

justice can be addressed. First of all, we can think of principles that reflect our intuitions about equal justice, indeed two principles have dominated the discussion on social justice: responsibility and compensation. Secondly, the sense in which justice encompasses equality is determined by the criterion of distribution which represents the best endorsement of equality. I suggest we start with a detailed inquiry into the principles of compensation and responsibility.

In Chapter 5 of Utilitarianism, J.S.Mill investigates the connection between justice and utility. Here we find one of the most persuasive and succinct accounts of the two dominant but conflicting principles of justice. Mill asks whether, in a co-operative industrial association, it is just or not that talent or skill should give a title to superior remuneration. On one side, Mill claims that it should not:

whoever does the best he can, deserves equally well, and ought not in justice to be put in a position of inferiority for no fault of his own. (J.S.Mill 1972, p.60)

On the other hand, Mill recognizes that it should:

society receives more from the more efficient labourer; that his service being more useful, society owes him a larger return for them; that a greater share of the joint result is actually his work, and not to allow his claim to it is a kind

of robbery. (J.S.Mill 1972, p.60)

According to Mill, justice has two sides to it, which are impossible to appease and harmonize. In order to break the deadlock, Mill's solution is to appeal to a third principle, namely, social utility.

I believe Mill has pinpointed, with characteristic clarity and precision, the key question at the centre of all debates on social justice. Nevertheless, I believe there is at least one alternative solution to the problem concerning social justice to the one favoured by Mill. Contrary to Mill's own views, I believe it is possible to bring the two sides of justice into harmony; this is the goal of the compatibilist theory of justice I will be defending in this chapter.

The two positions discussed by Mill can respectively be labelled the principle of compensation, and the principle of responsibility. Before attempting to provide an alternative answer to Mill's principle of social utility, it is important to explore the two moral intuitions behind the principles of compensation and responsibility.

Consider the following example. After a shipwreck, Ottavio and Anna drift on to a tropical island where the only edible type of food is coconuts. Let's assume that the two individuals set out searching for food, and to make their search more efficient they set out in different directions: Ottavio goes to the east, and Anna to the west.

At the end of the day they meet again. Ottavio has brought back eighteen coconuts, Anna only two. At this point Ottavio says: "I worked hard to find these eighteen coconuts, therefore I have a right to all eighteen of them". Anna replies "I worked hard to find my two coconuts, but I was unlucky to explore the western side of the island where there are very few coconut trees. Furthermore it is not my fault if I am only 5 feet tall, while you are seven feet tall, hence I cannot reach the coconuts high on the tree. Therefore we should divide the twenty coconuts equally, taking ten each". The argument put forward by Anna rests on the principle of compensation for ill-luck. Justice should compensate for brute bad luck, in this case the unfortunate choice of a direction or the fact of not being taller.

Consider now another example. Let's assume that two more individuals, Masetto and Elvira, are again stranded on a tropical island, where the same conditions as in the example of Ottavio and Anna apply. Desperate for food, Masetto and Elvira decide to look for coconuts by exploring different directions, hence Elvira goes north and Masetto south (the coconut trees are evenly spread between the north and south coasts of the island).

At the end of the day Elvira brings back eighteen coconuts and Masetto only two. Masetto says: "I worked as hard as I could all day, it is not my fault if I could only find two coconuts, therefore we should divide the total in half". Elvira replies: "I have worked hard to find these eighteen coconuts, never taking a rest. I climbed on the top of a tree to fetch a coconut and from

there I could see you sleeping under the shade. Throughout the day I have checked on you regularly and you were always sleeping in the shade. You are lazy and you want to eat the fruit of my labour. Therefore I am going to keep all my coconuts". The argument put forward by Elvira rests on the principle of responsibility. Justice should not compensate for inequalities when these are the result of voluntary acts for which we are directly responsible, in this case the choice to being lazy rather than productive.

I hope the two examples show that in determining the core principle of social justice, it is impossible to choose between responsibility and compensation, since both arguments are intuitively valid. Both Anna and Elvira are advancing a valid claim, hence a theory of justice that takes account of only one criterion at the expense of the other must be incomplete and unsatisfactory.

In order to show that the principles of responsibility and compensation have an intuitive appeal, I have made use of two hypothetical examples featuring Ottavio, Anna, Masetto and Elvira. Some readers may feel that the hypothetical examples I have used are hopelessly unrealistic, abstract and ultimately unconvincing, so that these examples cannot be trusted to validate our moral intuitions. Before going any further, I want to justify the nature of the above examples.

The purpose of the two examples was simply to add substance to the principles of responsibility and compensation, while emphasizing their basic moral intuitions. The advantage of abstract

examples is that they do not force us to think in terms of policy recommendations at an early stage in our argument. The two paradigmatic positions espoused by Anna and Elvira are both intuitively correct, hence they must be seen as two sides of the same coin. This is important because the starting point of our investigation on social justice must be the recognition that there is more overlap than fracture between the principles of responsibility and compensation.

In what follows, I want to consider how the principles of compensation and responsibility apply to the distribution of benefits and burdens. My argument will be divided in two parts. First, I will show that the principles of responsibility and distribution advance two radically opposed criteria of distribution, respectively choicism and anti-brute luck. Subsequently, I will put forward a solution that seeks to reconcile these two criteria.

6.II. Choicism and Anti-Brute Luck.

The criteria of choicism and anti-brute luck cover the full range of possible solutions to the problem of distribution. The core postulate behind these two arguments have been captured by Brian Barry in his "Chance, Choice, and Justice", indeed in this section I will be following closely Barry's exposition.

The anti-brute luck argument claims that "the victims of ill

luck should as far as possible be made as well off as those who are similarly placed in all respects other than having suffered this piece of bad luck." (Barry 1991, p.142). Clearly the anti-brute luck argument stems from the principle of compensation, namely, compensation for ill-luck.

On the other hand the choicist argument maintains that "social arrangements should be such that people finish up with the outcomes of their voluntary acts." (Barry 1991, p.142). Once again, it is easy to spot the correlation between this argument and the principle of responsibility, in fact the choicist argument tells us that ultimately we are responsible for the outcomes of our acts.

It is important to remember that the choicist and anti-brute luck arguments are conceptual paradigms¹¹². In other words, the choicist and anti-brute luck approaches are umbrella-categories covering a vast spectrum of positions. For example, the most extreme choicist position is given by Nozick in Anarchy, State, and Utopia, whose views can be summed up by the maxim "from each as they choose, to each as they are chosen". At the opposite pole, the extreme anti-brute luck position is given by Marx, whose views can be summed up by the maxim "From each according to their ability, to each according to their needs".

These two extreme positions are equally unpersuasive since they totally and unremittingly deny the opposite camp any say in

¹¹² - This adds justification to my choice of abstract and unrealistic examples in section 6.I; because of the theoretical nature of the choicist and anti-brute luck arguments, their respective core theses can only be captured via hypothetical thought experiments.

the matter. For example according to Nozick, voluntary exchange is the only valid criterion of distributive justice, indeed there is nothing unfair about the state of affairs whereby the only thing preventing Anna from starving to death is the generosity of Ottavio. On the other hand, according to Marx need is the only valid criterion of distribution, where need is directly proportional to brute luck. Yet between these two extremes there is a whole range of positions which allows for some degree of amalgamation between the choicist and anti-brute luck arguments.

In what follows I want to reject the two extreme positions of Nozick and Marx as non-starters, and concentrate on the various degrees in which the choicist and anti-brute luck arguments can overlap. For simplicity's sake, these positions can be depicted as points on a continuum line:



Figure 6.1

Positions A, B and C belong to the choicist camp, while positions D, E and F belong to the anti-brute luck camp. We have already met positions A and F, they are advocated respectively by Nozick and Marx. The other four positions (B, C, D and E) belong to either the choicist or anti-brute luck camp, although they all

recognize to some degree the validity of the opposing camp. In other words we are not facing an either/or question, instead the question here is one of proportion or degree of overlap.

Following Brian Barry's (1991) argument, although not always his terminology, I will refer to these four intermediate positions as follows:

- B - Extreme choicism
- C - Moderate choicism
- D - Moderate anti-brute luck
- E - Extreme anti-brute luck.

I have said before that choicism is an umbrella-category which embraces a number of positions. Thus while a choicist is someone "who wants to show that it is consistent with justice to give the principle of personal responsibility a lot of scope at the expense of the principle of compensation, while acknowledging the validity of both principles" (Barry 1991, p.144), we can distinguish between two separate positions, which I will call extreme choicism and moderate choicism.

Extreme and moderate choicism are distinguishable by the fact that the former is closer to Nozick's position while the latter is closer to the anti-brute luck argument. This means that extreme choicism believes that as long as everyone has access to the same resources, all subsequent inequalities can be defended as arising from voluntary choices; thus we find for example that extreme

choicists appeal to the institution of insurance as a justification for their approach¹¹³. On the other hand moderate choicism believes that although choice and responsibility are fundamental, the significance of choice in justifying unequal outcomes has been overstated; for example we are not always responsible for our choices, or alternatively, even if everyone has equal access to the same resources, these may in fact prove to be more favourable to one person rather than another. For these reasons moderate choicism renounces the insurance argument so dear to extreme choicism.

The same distinction that applies to the choicist camp, also applies to the anti-brute luck camp between extreme and moderate anti-brute luck arguments. Extreme anti-brute luck takes a strong deterministic line, arguing that no tastes, aspirations, or beliefs can ever be chosen or changed as a result of choice; since people cannot be held responsible for their tastes, aspirations, or beliefs, they cannot be held responsible for the outcomes of actions that flow from them. Moderate anti-brute luck abandons the deterministic stand, and instead settles for a case-by-case approach: some preferences, it is supposed, will give rise to personal responsibility, hence justifying subsequent inequalities, although many others will not.

So far I have argued that the two opposing camps of choicism and anti-brute luck hold a number of positions from extreme to moderate. I have also argued that a theory of justice must find a

¹¹³ - "Someone who loses everything in a fire can justly be left destitute because the option of insuring against the loss was available and not taken" (Barry 1991, p.145).

way of reconciling these two camps, since each camp is motivated by a valid intuition. In what follows, I will attempt to reconcile the principles of responsibility and compensation by accommodating choicism and anti-brute luck under one criterion of distribution, namely, the semi-choicist solution.

6.III. The Semi-Choicist Solution.

Of course managing the reconciliation between choicism and anti-brute luck is not as simple as it may appear, since there is a strong risk that one of the two camps will collapse under the weight of the other camp. The difficulty for a theory of justice, as Barry rightly points out, is not only to reconcile these two criteria, but to do so in such a way that one criterion does not collapse into the other criterion. That is to say, anti-brute luck should not swallow up choice, nor choice swallow up anti-brute luck.

If we follow our intuition that a theory of justice must be capable of reconciling compensation with responsibility, then we must investigate the possibility of striking a comfortable compromise between anti-brute luck and choicism. It is surprising to find that not many philosophers since J.S. Mill have attempted to reconcile the principles of compensation and responsibility in a theory of justice. So far the best attempt to tread this middle road is Barry's account of the semi-choicist solution. In what

follows I will first analyze Barry's semi-choicist solution. Subsequently, I will argue that in order to make the semi-choicist solution as strong as possible, it is necessary to supplement Barry's general argument with a supplementary criterion.

Barry's own solution to the problem of reconciling choicism with anti-brute luck is what he calls the semi-choicist position. His starting point is to accept the anti-brute luck premise that it is legitimate to look behind choices to preferences before establishing personal responsibility for outcomes: "People are, indeed, responsible for the outcomes of actions only if they are also responsible for the preferences from which those actions flow" (Barry 1991, p.156). Yet beyond this point Barry's takes a different line from the anti-brute luck position. While the anti-brute luck criterion makes responsibility for preference depend upon their being subject to choice, Barry suggests an original account of what it means to be responsible for one's own preferences:

The alternative I propose is that people are responsible for their preferences whenever they are content with them. How these preferences originated is irrelevant, and the ease with which they could be changed is relevant only in this way: that we would have to question the sincerity of your claim not to want to have the preferences you actually do have if it were easy for you to change them.

I shall call this view 'semi-choicism', since it abandons

the choicist resistance to any inquiry into responsibility for preferences, but rejects the anti-choicist criterion for responsibility for preferences. (Barry 1991, p.156; emphasis added)

According to Barry, the whole idea of choosing preferences is misguided. Thus instead of making responsibility for preference depend upon their being subject to choice, the semi-choicist solution suggests that one is responsible for one's own preferences if one is content with one's preferences.

In order to appreciate the strength of Barry's solution, I suggest we consider the following example. Leporello and Giovanni are two junior lecturers in tenure track positions. At the end of their three-year temporary appointment, we find that Leporello has published a number of articles in important journals, and thus is given tenure. On the other hand Giovanni has not published a single article and is not given the tenure position. Is this fair?

We know that Giovanni is a keen bridge player, and everyday after his teaching is over he goes to his bridge club. Considering this, the anti-brute luck criterion would tell us that there is nothing unfair about the awarding of tenure to Leporello but not to Giovanni, since Leporello has chosen to pursue his academic career during his free time, while Giovanni has chosen to perfect his bridge playing. To the extent that they have chosen their preferences, Leporello and Giovanni are responsible for their actions.

It would appear that this first solution works, certainly in the case of the two lecturers it intuitively seems to give the right answer. Yet this idea of choices of preferences has a major fallacy, namely, it assumes that we have complete control over our preferences, while in fact it has been argued that that is not the case.

Of course one of the problem we face is that the concept of preferences defies a precise definition, therefore it is difficult to say anything conclusive about preferences. Yet if we assume that preferences refer to our tastes, aspirations and beliefs, it is not always the case that we have control over our preferences. For example, to what extent do I have control over the fact that I like strawberries more than apples, or that I have a certain sexual inclination, or that I hold certain religious beliefs¹¹⁴?

If we question our ability to modify our preferences, then we must allow for cases where a person is not responsible even if that person has certain preferences. This leads us to the second possible solution, namely, the semi-choicist solution. Consider the following example. A person may suffer from certain preferences, such as smoking, drinking, drug abuse or eating disorders. Although these are the result of certain preferences, we can argue that to the extent that this person genuinely wants to change these preferences, he or she should not be considered responsible.

¹¹⁴ - Religious beliefs are usually the result of a certain upbringing, which implies that a person has very little control over his or her religious beliefs.

The problem of dealing with eating disorders is particularly revealing. As Naomi Wolf (1990) argues, the beauty myth may drive a woman (say Zerlina) to become bulimic or anorexic. Yes, it is her preference to become 'beautiful', furthermore it was her choice to follow a diet, but she is certainly not content with the state she is in now. Hence we could argue that Zerlina should be offered psychotherapy on the NHS.

Barry's semi-choicist solution does a better job of dealing with Zerlina's case. Barry suggests that people are responsible for their preferences whenever they are content with them, and not if the preferences are the subject of choice. Thus in the case of Zerlina, we have to ask ourselves if Giovanni is content with the choice of her preferences. I think it is fair to assume that she would rather not have been anorexic or bulimic, even if it was her choice to follow a diet. Zerlina's case shows that one is not always responsible for the outcome of one's actions even if those follow from chose preferences.

I believe Zerlina's example shows the limits of the anti-brute luck argument, and the advantages of the semi-choicist solution. Both theories would agree that Giovanni should not be given tenure, although they give different reasons: according to the anti-brute luck argument, the problem is that Giovanni chose his preference to play bridge, while according to the semi-choicist argument, the problem is that he was content to play bridge. It follows that in cases such as Zerlina, the two theories give different responses. According to the anti-brute luck argument, Zerlina has chosen her

preference of following a diet, hence she should not claim for compensation, while according to the semi-choicist solution Zerlina should be compensated. In cases such as Zerlina's, the solution given by the semi-choicist argument appears to be considerably more convincing than the anti-brute luck argument.

While Barry's semi-choicist solution, based on the criterion of content for one's preferences, is the best attempt to reconcile compensation with responsibility, it can be argued that there is still room for improvement.

First of all, although in the case of Zerlina's eating disorder Barry's criterion of content seems to give the right response, there appears to be a complication with Barry's semi-choicist solution, namely, it is not clear whether the criterion of content applies to the preferences as such, or to the consequences of acting on one's preferences. Let's assume that I go to a party with a group of vodka-loving Russian friends: does content refer to my preference to drink vodka rather than orange juice, or does it refer to the consequences of my actions (the fact that I am not content with the hangover resulting from drinking two bottles of vodka in one night).

Secondly, it seems to me that Barry's criterion of being content with one's preferences is not sufficiently robust to stand as the sole foundation for the semi-choicist solution¹¹⁵. In fact

¹¹⁵ - At the end of his article on "Chance, Choice, and Justice", Barry points out that his analysis of chance and choice is only a brief venture in this subject area, and that "More work needs to be done to refine the analysis and to apply it" (Barry 1991, p.158), indeed Barry has promised his readers a more thorough

while in most cases we can identify responsibility by equating one's preferences with one's contentment, there are some cases when one is not responsible for one's preferences even if one is content with them, as well as cases when one is responsible for one's preferences even if one is not content with them.

I am referring to cases when one acts upon a preference in ignorance of the possible repercussions it may have, hence one is content of one's preference only because one did not know better. In order for the semi-choicist solution to take these cases into account, it is necessary to expand the criteria of responsibility for one's preferences. Thus I believe that in order to establish one's responsibility one must (a) look at the preferences and ask whether one is content with them, and (b) establish whether one is accountable for the consequences of one's preferences. I will refer to the latter condition as the criterion of liability.

The limits of Barry's semi-choicist solution can be exposed by considering another type of example. There are two patients in a NHS hospital, where (needless to say) resources are scarce. Both patients suffer from smoking-related illnesses, and require a lung transplant. Patient A is 56 years old, and started smoking in 1954 when he was 16. Patient B is 26 and started smoking in 1984 when she was 16. If the hospital has the opportunity of carrying out one transplant immediately, who should be the lucky recipient? We know

examination in a forthcoming volume of his Treatise on Social Justice. What follows is not meant to be a critique of Barry's theory, but simply a contribution to the analysis of chance and choice.

that the cause of their illness is smoking, and we must assume that smoking is a preference. Furthermore if we examine the preferences of the two patients, we find that they are both content to be smokers. Should we deduce from this that there is nothing distinguishing patients A and B, and therefore some form of lottery is the best form of resource allocation?

I want to argue that there is a major difference between patient A and patient B that is not detected by Barry's criterion of content. The difference is that when patient A started smoking in 1954, he could not have known of the health hazard associated with the habit, while patient B, who started smoking in the 1980s, must have been particularly stubborn to neglect the anti-smoking campaigns (not to mention the Government's health warnings against smoking printed in bold letters on every packet of cigarettes), hence patient B should have known better than to become a smoker. The case of the two patients illustrates what can be called the liability question, since the evaluation of the responsibility of the two patients over their preferences takes into account whether they could-have-known-better.

The example of the two NHS patients shows that Barry's criterion of content is not sufficiently comprehensive to take into account all possible cases. In particular, the criterion of content is incapable of identifying cases where someone is not-responsible even if the person is content with his/her preferences. This is why I feel that Barry's criterion of content should be supplemented by the criterion of liability, in other words when assessing people's

preferences we should consider both the question of content and the question of not-knowing-better.

The idea that responsibility is related to the fact that "one should have known better" is highlighted by Scanlon (1988b) in his account of justice and choice. Scanlon considers the example of the removal and disposing of some hazardous waste. Although the authorities did all they could to warn and protect them (by building a high fence around the excavation site, placing large signs warning of the danger, arranging for the removal and transportation to take place at night, etc.), some chemicals were released into the air. Sometime later, two persons suffered from lung damage. Person A was unfortunate to be particularly sensitive or even genetically predisposed, and suffered as a result of the chemicals released into the air. Person B heard the warnings but did not take them seriously, and her curiosity led her to climb the fence and inspect the site herself. According to Scanlon, Person B "bears the responsibility of her own injury By choosing, in the face of all warnings, to go to the excavation site, she laid down her right to complain of the harm she suffered as a result" (Scanlon 1988b, p.192).

Let's recapitulate the argument so far. The primary objective of this chapter is to show that the principles of compensation and responsibility are both equally valid, hence the problem liberal political philosophy faces is not choosing one or the other, but instead reconciling these two principles. By looking at the problem of establishing the best criterion of distribution, I have argued

in favour of a solution which successfully reconciles the criteria of distribution implied by the principles of compensation and responsibility.

Having established a criterion of distribution that appears to successfully reconcile the principles of responsibility and compensation, it is now necessary to situate this criterion within the wider scope of a theory of justice. Fortunately this project is less daunting than it may appear, since there is already a theory of justice that portrays the characteristics of compatibility between the principles of responsibility and compensation. In what follows, I will argue that the semi-choicist solution is not adverse to Rawls's theory of justice, indeed I believe that Rawls's second principles of justice exemplifies aspects of compensation as well as responsibility.

6.IV. The Principle of Responsibility and Justice as Fairness.

The proposed attempt to reconciling the compatibilist theory of justice presented in parts 6.II and 6.III, with Rawls's theory of justice as fairness, may appear (to say the least) extravagant. Indeed this is not surprising. Considering Rawls's well publicised severe rejection of desert¹¹⁶, it may be argued that there is no place for the principle of responsibility in Rawls's theory of

¹¹⁶ - See Rawls (1971), pp.103-4. For a critical account of Rawls's anti-desert line, see Sher (1987), Ch.2.

justice.

Considering this, before proceeding any further I feel I ought to spell out as clearly as possible how I intend to utilize Rawls's theory of justice, and what I mean by claiming that the reconciliation between the principles of compensation and responsibility can be accommodated in the framework of Rawls's two principles of justice.

I will start by stating what I am *not* advocating in this part of the chapter. First of all, I am not making the claim that it was Rawls's ambition to reconcile the two opposing principles within one theory. Clearly that is not the case, and we must accept it. It follows that I am not presenting a new interpretation of Rawls's theory of justice that pretends to do a better job of explaining the intentions of the author.

My more modest claim is simply the following. Starting from the fact that Rawls's failure to find room for the principle of desert in his theory of justice is arguably one of the biggest weaknesses with Rawls's theory¹¹⁷, I believe it is necessary to work on the basis of Rawls's theory in order to redress this lacuna. Since the idea of desert is grounded on the concept of responsibility (I will return to this claim later), by showing that there is evidence of both responsibility and compensation in Rawls's second principle of justice I will argue that Rawls's

¹¹⁷ - Brian Barry points out that the response to Rawls's strong anti-desert line since 1971 has been overwhelmingly negative: "I cannot as a matter of fact think of anything that has been written defending Rawls on desert..." Barry (1990c), p.lvii.

theory of justice need not be as radically anti-desert as Rawls claims. Another way of expressing this point is by saying that the reconciliation between responsibility and compensation is not alien to Rawls's project, even though such reconciliation was not Rawls's intention.

I should also add that my attempt to squeeze the semi-choicist solution within the general framework of Rawls's theory of justice is not a violation of Rawls's theory. Indeed Rawls would be the first to admit that the two principles of justice are not meant to be exact rules but only general indications for institutions. It is important to recall that after giving the final statement of the two principles of justice, Rawls makes the following claim:

By way of comment, these principles and priority rules are no doubt incomplete. Other modifications will surely have to be made, but I shall not further complicate the statement of the principles. (Rawls 1971, p.303)

The detailed argument presented in the first part of this chapter concerning the semi-choicist solution is nothing more than an attempt to modify and specify (although perhaps also complicate) Rawls's two principles of justice as fairness.

Let's recall Rawls's two principles of justice as fairness:

First Principle - Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a

similar system of liberty for all.

Second Principle - Social and economic inequalities are to be arranged so that they are both:

- (a) To the greatest benefit of the least advantaged (difference principle)
- (b) Attached to offices and positions open to all under conditions of fair equality of opportunity (fair opportunity)

Of the two principles of fairness, the difference principle has attracted the greatest interest over the years. While it cannot be denied that the difference principle is the heart and soul of Rawls's theory of justice¹¹⁸, it is also important to remember that the difference principle is only one part of the second principles of justice, that the two principles of justice follow a lexicographical ordering, and that the difference principle is preceded by the other principles¹¹⁹.

On the basis of the role and place of the difference principle in Rawls's account of the two principles of fairness, I want to put forward the following suggestion: the principles of responsibility and compensation can be reconciled with Rawls's theory of justice

¹¹⁸ - As Barry point out, the difference principle is not a second-best principle of justice, instead the difference principle is "what justice really demands" (Barry 1989b, p.398). It is also important to emphasize that the difference principle constitutes the core aspect of what Rawls calls the general conception of justice.

¹¹⁹ - Not only the first principle comes prior to the second principle, but within the second principle "fair opportunity is prior to the difference principle" Rawls (1971), p.303.

to the extent that the principle of responsibility is enshrined in the principle of fair opportunity, while the principle of compensation is enshrined in the difference principle.

The equation between the principle of responsibility and the principle of fair opportunity becomes clearer if we ask ourselves why Rawls incorporates a principle of fair opportunity in his theory of justice, and why this principle has priority over the difference principle. If Rawls was exclusively interested in the principle of compensation¹²⁰, he could have either neglected the principle of fair opportunity, or at least have given lexical priority to the difference principle over the principle of fair opportunity. But as we have seen that is not the case. The principle of fair opportunity is an important part of Rawls's theory of justice, even if it has been generally neglected by commentators who have been more keen to defend or accuse the difference principle. I believe the principle of fair opportunity has important similarities with the principle of responsibility.

In order to see the affinity between responsibility and opportunity, it is necessary to analyze the concept of opportunity in more detail. First of all, it is important to stress that there is more to opportunity than the instrumental value, or in other words opportunity is not valued simply for its potential

¹²⁰ - In Rawls's terminology, this would be compensation for the unfortunate lack of bargaining advantages resulting from the social and natural lottery.

consequences. Scanlon (1988b) argues that the value of choice¹²¹ can be appreciated from three angles: instrumental (we want future events to conform to our preferences), demonstrative (the act of choosing has a special meaning, for example choosing a gift for a person we love), or symbolic (choosing may be a way of demonstrating competence, for example I may choose a wine at a restaurant in order to avoid the embarrassment of admitting my ignorance of wines).

Yet whether a choice or opportunity is regarded for its instrumental, demonstrative or symbolic value, in all three cases the choice or opportunity necessarily assumes a sense of responsibility over one's action. It follows that equality of opportunity over offices and positions is important because of the value persons attach to opportunities. Equality of opportunity is to be valued because of what we can achieve once we have been given an opportunity, and we are at least partially responsible for what we achieve. In other words, equality of opportunity is important if and only if we are also responsible (to some extent) for our actions. If the principle of responsibility plays no part in the proceedings, then equality of opportunity is redundant.

It is also interesting to note that fair opportunity comes prior to the difference principle. I believe this can be explained by the fact that if it was the other way around, and the difference principle came prior to fair opportunity, the latter could not be

¹²¹ - Scanlon uses the term 'choice' to mean both voluntary action and opportunity.

implemented. Indeed this would be a classic case of the principle of compensation swallowing the principle of responsibility. The difference principle is a principle of inequality, while fair opportunity is a principle of equality. It is more difficult (if not impossible) to justify a shift from acceptable inequality to equality than vice-versa, hence if the difference principle was prior to fair opportunity, the latter would disappear.

Before turning to the difference principle and the principle of compensation, I want to consider two issues concerning the question of desert in Rawls's theory of justice. In fact by arguing that the principle of responsibility has room in Rawls's theory of justice, it may seem that I am re-introducing desert in a theory which was explicitly meant to be anti-desert.

First of all, I believe that Rawls's theory of justice has more to lose than gain from refuting the principle of responsibility. Indeed denying responsibility implies endorsing a form of determinism that is irreconcilable with other aspects of Rawls's theory, for example the fact that the persons in the original position are rational maximisers (Rawls 1971, pp.142-150), that the principles of justice apply to primary goods or basic resources rather than welfare (Rawls 1971, pp.92-95), and that we have the moral powers to formulate, pursue and revise our conceptions of the good (Rawls 1980).

Secondly, I want to come back to my claim that the idea of

desert is grounded on the concept of responsibility¹²². The concept of desert is very complex, and my comments on desert and responsibility do not pretend to be original or in any way revealing. Indeed my comments on desert do not even pertain to all the meanings of desert or to all instances where this concept may apply. By desert here I mean the idea of desert as merit, that is to say, the idea that a person deserves something by virtue of some standard of excellence. In the words of Sher:

Merit spans both the moral and the nonmoral realms....The morally meritorious include both people who perform single transcendent acts of heroism or sacrifice and persons whose generosity or compassion is woven through their lives. The nonmorally meritorious include athletes who run faster than others, scientists who discover cures for deadly diseases, and jobs applicants who score highest on qualifying exams. (Sher 1987, p.109)

The reason why I feel justified to concentrate exclusively on desert as merit is because I feel that this is what Rawls was thinking when he rejected desert. In other words, Rawls's theory of justice is anti-merit rather than anti-desert.

Although the idea of desert as merit does not necessarily

¹²² - Recently there has been renewed interest on the question of desert and responsibility in contemporary liberal philosophy. See Scheffler (1992) and Ripstein (1994).

require a rejection of determinism¹²³, I believe desert does imply a minimal level of responsibility. That is because unless the concept of responsibility is part of desert, it becomes impossible to distinguish the idea of desert from the idea of luck¹²⁴. Generally, claims of desert have a common structure, being constructed by three elements: Person A deserves B in virtue of C. This implies that Person A must do something in order to deserve B, that is to say, something must change before Person A deserves B, and that change must be brought about by A. Luck cannot be the requirement that makes person A deserve B, since by being lucky person A has not changed anything. If I find a twenty-pounds note in the street, I cannot claim to deserve it any more than any other Londoner, instead I am simply lucky to be at the right place at the right time. Luck does not require effort, hence it cannot be the basis for desert.

If my claim that there is an affinity between Rawls's second principle of justice and the principle of responsibility is right, and that the principle of responsibility is an important assumption for any argument of desert, then it becomes possible to placate the familiar critique of Rawls's violent rejection of desert without abandoning Rawls's two principles of justice as fairness.

¹²³ - See Barry (1990c), p.lxi.

¹²⁴ - For an interesting discussion on responsibility and luck, see Nagel's "Moral Luck" in Nagel (1979).

6.V. Cohen's Critique of the Difference Principle.

Although the equation between the principle of responsibility and Rawls's principle of fair opportunity may be a controversial claim, the equation between the principle of compensation and the difference principle would appear to be unquestionably sound. The principle of compensation claims that the victims of ill luck should as far as possible be made as well off as others who have not suffered this piece of bad luck, and the difference principle states that social and economic inequalities are to be arranged so that they are to the greatest benefit of the least advantaged. To the extent that the victims of bad luck are the least advantaged, the difference principle acts as a principle of compensation.

Yet recently the validity of the difference principle as a principle of compensation has been questioned. In his Tanner Lecture "Incentives, Inequality and Community", G.A. Cohen gives a critical evaluation of Rawls's difference principle, arguing that as a principle of social justice the difference principle is impaired by some serious limitations. In what follows I will evaluate and reject Cohen's critique.

Cohen's major critique of Rawls's difference principle is that inequalities associated with incentives satisfy the difference principle. The incentives argument claims that because of the trickle-down effect, the worse off are better off if the well off are even better off:

high levels of income cause unusually productive people to produce more than they otherwise would; and, as a result of the incentives enjoyed by those at the top, the people who end up at the bottom are better off than they would be in a more equal society. (Cohen 1992, p.265)

The reason why Cohen rejects the incentive argument is because it fails what he calls the interpersonal test, indeed Cohen claims that a policy argument provides a comprehensive justification only if it passes the interpersonal test. Cohen's interpersonal test checks for variations with respect to who is speaking and/or who is listening when the argument is presented. Thus a policy argument passes the interpersonal test if it can be shown that it could be uttered by any member of society to any other member. Clearly this is not the case of the incentive argument; the incentive argument is an argument made by the rich to fit their interests, although it is disguised under the impartial cloth of sociology and economics.

It should be pointed out that there are interesting and important similarities between Cohen's interpersonal test, and Scanlon's impartiality-based contractualism. The importance of this similitude lies in the fact that it uncovers a contradiction in Rawls's argument for the difference principle. In fact if the interpersonal test checks for impartiality, and the incentive argument fails the interpersonal test, then it follows that by endorsing the incentive argument Rawls's difference principle fails the impartiality test.

In Chapter 3 it was shown that according to Scanlon's understanding of impartiality, we are moved by a desire to justify one's actions to others on grounds they could not reasonably reject. Cohen's interpersonal test seems to follow from the same idea, indeed any policy argument that passes Cohen's interpersonal test would also pass Scanlon's impartiality test. Nevertheless there is a small but significant difference between Cohen and Scanlon, indeed following Scanlon's terminology we could rephrase Cohen's interpersonal test in terms of justifying one's actions to others on grounds oneself could not reject. That is to say, it is not sufficient to justify my actions to you on grounds you could not reject, but I should also justify my actions to you on grounds I could not reject. An example should make this clear.

Cohen points out that the incentive argument does not serve as a justification of inequality on the lips of the talented rich because it would not pass the interpersonal test, indeed the incentive argument is uttered by the rich as part of an impersonal discourse of sociology or economics, although in fact this argument takes the form of a discourse grounded on willing choices. For example Cohen argues that in the case of top executives, they are willing to work hard only if they pay smaller taxes. The key word here is 'willingness'; top executives "could not justify the behaviour the argument ascribes to them" (Cohen 1992, p.282; emphasis in original), since willingness does not suffice as a justification.

If we return to Cohen's critique of Rawls, we can see now that

Rawls's difference principle endorses the incentive argument, although the idea of incentives fails the interpersonal test, hence the difference principle is accused of lacking impartiality. The core argument in Cohen's article is that, when true to itself, Rawlsian justice condemns incentives to people of talent, "and that no society whose members are *themselves* unambivalently committed to the difference principle need use special incentives to motivate talented producers" (Cohen 1992, p.310; emphasis in original).

Although Cohen is critical of the difference principle, it is not his intention to undermine this principle tout court, indeed at the beginning of the lecture Cohen reveals his acceptance of the principle: "For my part, I accept the difference principle but I question its application in defense of special money incentives to talented people" (Cohen 1992, p.268). The aim of Cohen's argument is to defend the following claim: the difference principle is either valid as a principle of public policy, although not as a fundamental principle of justice; or it applies to a society where people are *themselves* moved by an ethos or culture of justice, in which case the difference principle cannot be implemented by a government.

In order to support his claim, Cohen maintains that there are two readings of the difference principle: a strict and a lax reading:

in its strict reading, it counts inequalities as necessary only when they are, strictly, necessary, necessary, that is,

apart from people's chosen intentions. In its *lax* reading, it countenances intention-relative necessities as well. So, for example, if an inequality is needed to make the badly off better off but only given that talented producers operate as self-interested market maximizers, then that inequality is endorsed by the *lax*, but not by the strict, reading of the difference principle. (Cohen 1992, p.311; emphasis in original)

According to Cohen, both readings of the difference principle are present in Rawls's theory. The first is 'intention-independent', and it refers to Rawls's account of a well-ordered society. The second is 'intention-sensitive', and it reflects Rawls's endorsement of incentives¹²⁵.

Cohen argues that if we defend the *lax* difference principle along the lines of a reasonable compromise between self-interest and service, it follows that the *lax* difference principle "is at best an imperfect proxy for a just balance, and not, what it is supposed to be, a fundamental principle of justice" (Cohen 1992, p.315). Cohen concludes: "I have not rejected the difference principle in its *lax* reading as a principle of public policy: I do

¹²⁵ - These two readings of the difference principle reflect an underlying tension between opposing conceptions of social relationships. Thus the intention-sensitive reading of the difference principle reflects a bargaining conception of social relationships, while an intention-independent reading reflects a community conception of social relationships. Cohen's analysis here seems to be heavily influenced by Barry's (1989b) argument that Rawls embraces two opposing theories of justice.

not doubt that there are contexts where it is right to apply it. What I have questioned is its description as a principle of (basic) justice..." (Cohen 1992, p.328).

It follows that we are left with the strict difference principle, although Cohen claims that government by itself cannot implement it:

For the strict difference principle to prevail, there needs to be an ethos informed by the principle in society at large. Therefore, a society (as opposed to its government) does not qualify as committed to the difference principle unless it is indeed informed by a certain ethos, or culture of justice. Ethos are, of course, beyond the immediate control of legislation, but I believe that a just society is normally impossible without one, and Rawls himself requires that there be a nurturance and cultivation of appropriate attitude in the just society that he describes. (Cohen 1992, pp. 315-6)

I want to argue, contra Cohen, that Rawls's difference principle can act as a fundamental principle of justice. In order to justify this claim, I will argue that Cohen's attempt to diminish the importance of the difference principle, in its strict interpretation, by appealing to a society's ethos or culture, is both misleading and potentially harmful.

To state the point briefly, Cohen's appeal to a society's ethos is misleading because it treats a society's culture or ethos

as preceding the fundamental principles of justice, while I believe it would be more accurate to portray the relationship between these two as symbiotic.

It is interesting, and indicative, that Cohen never addresses the question of where a society's ethos or culture come from. Of course an adequate answer to this question moves us on to metaphysical grounds, and I for one don't have a satisfactory answer to my own question. Nevertheless it seems to me to be intuitively plausible that principles of justice have some effect on our ethos or culture, if not immediately at least in the long run. I am sure G.A. Cohen would agree with me that a society's ethos and culture is not static, but progressive, or at any rate open to mutation and change. Accepting this premise is only a small step away from the assertion that the choice of a fundamental principle of justice can have an impact in the shaping of a society's ethos or culture. I believe that Rawls has in mind something along these lines when he refers to the educational role of constitutionalism in general and justice as fairness in particular:

Some have thought that if a people is truly democratic in spirit, a constitution with a bill of rights is unnecessary; while if a people is not democratic, such a constitution cannot make it so. But this view overlooks the possibility that certain features of a political conception importantly affect the political sociology of the basic institutions that

realize it. More exactly, we must consider how that sociology may be affected by the educational role of a political conception of justice such as justice as fairness with its fundamental ideas of person and society. (John Rawls MS, p.107; emphasis added)

I can't see any reason why the educational argument presented by Rawls concerning constitutionalism and justice as fairness cannot also apply to the strict interpretation of the difference principle. In conclusion, although I am not denying that there is a close affinity between a society's ethos or culture, and its fundamental principle of justice, I don't see why we should assume, as Cohen does, that the former necessarily precedes the latter.

Apart from being misleading, I believe Cohen's argument based on a society's ethos and culture can also be potentially harmful. This is so for two separate reasons. First, because it opens a door to communitarianism, and secondly because it comes dangerously close to (unfairly) dismissing some societies as doomed.

This is not the place to discuss in any depth the question of communitarianism. What I intend to do instead is to focus on one specific weakness of communitarianism, namely, the arbitrariness of the concept of community from an ethical standpoint. By claiming that a society's ethos or culture precedes the difference principle as a fundamental principle of justice, Cohen is in fact arguing that a society has an independent ethos. It seems to me that Cohen is making an ontological assertion analogous to the communitarian

claim that the key to ethics lies in the practice of the community.

Yet apart from the fact that there does not seem to be a common and coherent communitarian position on the meaning of the term 'community'¹²⁶, communitarians always assume the superiority of the community over its composing parts, or assert the moral superiority of holism over individualism, without questioning the fact that its republican motivation can easily collapse into an apology for conservatism¹²⁷ or communism, both of which abhor conflict, diversity and pluralism¹²⁸. I believe the same criticism applies to Cohen's notion of a society's ethos or culture. While I am not denying that it is difficult to question the practice of a society's ethos and culture, it can be done. As Kymlicka rightly points out:

It may not be easy to question deeply held beliefs about the good, but the history of the women's movement, for example, shows that people can question and reject even the most deeply entrenched sexual, economic and family roles. We are not

¹²⁶ - As Gardbaum (1992) rightly points out, there are three separate debates involving three separate claims of community, namely: antiatomism, metaethical communitarianism; and strong communitarianism.

¹²⁷ - Roger Scruton claims that "conservatism presupposes the existence of a social organism. Its politics is concerned with sustaining the life of that organism, through sickness and health, change and decay" Scruton (2nd.ed.1984), p.25, quoted in Gardbaum (1992), p.722.

¹²⁸ - For the desirability of endorsing conflict as the key moment in a moral theory, see Hampshire (1983). See also Chapter 1 above.

trapped by our present attachments, incapable of judging the worth of the goals we inherited or ourselves chose earlier.
(Kymlicka 1993, p.371)

All I am suggesting is that a society's ethos or culture does not prevent us from adopting the difference principle as a fundamental principle of justice, even if the former is opposed to the latter.

Another problem with Cohen's approach is that there seems to be no hope for all those societies that don't hold the ethos or culture recommended by Cohen to grow to accept the difference principle as a basic principle of justice. In other words, if a society does not embrace an egalitarian ethos, does it mean that it will never embrace an egalitarian principle of justice? Take for example South Africa. The state of South Africa was until recently unquestionably unjust, and if the ruling class held any ethos or culture it is certainly not one compatible to Rawls's difference principle. Does it mean that South Africa had no hope of ever changing its ethos? After all, if Germany could go from national socialism to social democracy in the space of fifty years, why can't South Africa, Rowanda etc.?

Determining whether a society's ethos comes prior to its fundamental principles of justice, or vice-versa, may appear to be like solving the chicken-and-the-egg dilemma, but in fact this is not so. The question is whether changes are the result of internal or external pressures. Contrary to the view of Cohen, I believe

that we should not shut the door to the former solution, hence the difference principle should be given at least the benefit of the doubt as a motivating force for changing a society's ethos.

To recapitulate, in this part of the chapter I have argued that it is both misleading and potentially harmful to diminish the importance of the difference principle, in its strict interpretation, by appealing to a society's ethos or culture as a preliminary assumption. Therefore it seems to me that as a principle of compensation Rawls's difference principle (in its strict interpretation) is a valid and strong candidate. In what follows, I want to argue that the idea of reconciling the principles of compensation and responsibility under a theory of justice is warranted by the contractualist theory of reasonable agreement.

6.VI. Justice and Reasonable Agreement.

There are two key aspects to the theory of reasonable agreement. First, the moral motivation at the root of this theory is derived from the idea of agreement, not the concept of reasonableness. Second, the concept of reasonableness, defined by the maxim 'not asking for too much', provides the rationale that illuminates our intuitions. In what follows, I will show that these two aspects point to the necessity to endorse a compatibilist theory of justice.

Let's start with the *idea of agreement*. Agreements are the stuff of which social contracts are made. A contractarian political theory is, by definition, grounded on the notion of agreement. The agreements do not refer simply to the mechanics of accord between rational and consenting adults, although this is an important aspect of contractualism, instead it is the *idea of an agreement* that is fundamental to a social contract. By the *idea of agreement*, I mean that the parties seeking agreement are willing to compromise or revise their conception of the good in order to make room for others. The *idea of agreement* is important for a contractualist political theory since it enables us to distinguish legitimate from illegitimate interests or preferences; one's preference is acceptable if one is willing to compromise it, or in some way revise it, in order to seek agreement with others.

It seems to me that the *idea of agreement* justifies a commitment to both the principles of responsibility and compensation. First of all, the *idea of agreement* necessarily assumes that the parties are responsible for their actions, hence the principle of responsibility must be a component of a theory of justice grounded on reasonable agreement.

As I argued in Chapter 3, the *idea of an agreement* requires the conditions of critical reflection and rational self-governance. These two conditions are indispensable if the idea that someone is motivated by a desire to justify one's actions to others on grounds they could not reasonably reject is at all intelligible. In fact the concept of justification is significant if we are responsible

or accountable for our intentions and decisions. As Scanlon points out:

insofar as these intentions and decisions are ours, it is appropriate to ask us to justify or explain them Moral criticisms and moral arguments, on the contractualist view, consist in the exchange of such requests and justifications.

(Scanlon 1988b, p.171; emphasis in original)

While the idea of agreement justifies a commitment to responsibility and choicism, the principle of responsibility is not endorsed at the expense of the principle of compensation. In other words acknowledging that the principle of responsibility plays a part in the compatibilist theory of justice does not mean that choicism is the only criterion of distribution. Instead, it would be more accurate to say that the theory of reasonable agreement subscribes to the principle of responsibility only after the conditions for compensation have been fulfilled.

This can be shown if we recall that what is important about responsibility is the fact that people attach a value to opportunity and choice, whether this value is intrinsic or instrumental. It follows that a contractualist theory of reasonable agreement is keen to ensure that someone's value of choice will not undermine the value of choice of others (Scanlon 1988b, pp.186-7).

This is why a compatibilist theory of justice does not make an appeal to a preinstitutional notion of desert. Instead, as Scanlon

points out:

The only notions of desert which it recognizes are internal to institutions and dependent upon a prior notion of justice: if institutions are just then people deserve the rewards and punishments which those institutions assign them. (Scanlon 1988b, p.188)

In other words, unless desert is constrained within an institutional framework there is a risk of some people being deprived of their value of choice. This would be unjust. It is because people attach great importance on having outcomes depend on their choices that desert is not a preinstitutional notion.

The appeal of the principle of compensation is that it makes it possible for people to be responsible for their actions, and therefore to value their choices and opportunities. One can only be responsible for one's choices if the choices are made on the basis of appropriate conditions of equality of opportunity, and compensation ensures that such conditions are met.

Clearly the moral motivation defined by the *idea of agreement* is closely related to the concept of impartiality. Seeking an agreement with others and being willing to revise one's conception of the good captures the intuition, integral to the concept of impartiality, that the well-being of all matters intrinsically, or that all interests ought to be given equal weight. By respecting the *idea of agreement*, a theory of justice that reconciles the

principles of compensation and responsibility is congenial to the moral sanction of impartiality.

Having argued that the idea of agreement supports a reconciliation between the principles of compensation and responsibility, I now want to show the advantage for a compatibilist theory of justice of conceiving the idea of agreement in terms of a reasonable agreement.

In Chapter 3, I argued that in the theory of reasonable agreement, the concept of reasonableness should not replace the moment of agreement as the fulcrum of the contractarian moment, instead its role is to supplement our moral motivations. In other words, the concept of reasonableness is the rationale which qualifies our moral motivations, where the idea of agreement is the basis of our moral motivation.

I have also argued that the concept of reasonableness can be defined by the maxim 'not asking for too much'. Starting from the assumption that all demands are demands for something that is in limited supply, we can conclude that a demand is considered reasonable if it does not ask for too much of scarce resources compared to other demands for the same resource. The fact that everyone (including those who are not yet born) has an equal right to make a claim for these resources is an indication of the affinity between the concepts of reasonableness and impartiality. From the definition of reasonableness as 'not asking for too much', we can infer the significance of the notion of reasonable agreement, namely, an agreement is reasonable when no-one seeking

agreement is asking for too much of scarce resources.

The relation between reasonable agreement, the concept of reasonableness, and the compatibilist theory of justice is the following. The maxim 'not asking for too much' sets up a test for the contracting parties. This test will tell us if the parties are indeed seeking an impartial agreement¹²⁹, hence if their demands are legitimate. In other words, in order to establish whether the contracting parties are making reasonable demands, the parties involved must ask themselves if their demands are 'asking for too much' compared to the other parties.

To appreciate how this relates to the compatibilist theory of justice, we must start from the assumption that the principles of compensation and responsibility represent respectively the interests of those in need and those who merit. It follows that a theory of justice that is grounded exclusively on either the principle of compensation, or the principle of responsibility, is inevitably asking for too much from one of those two groups of people. Those who stand to gain most from the endorsement of one principle must ask themselves if they are asking for too much from others who are set to be net losers if that principle is adopted. For example if a theory of justice endorses only the principle of compensation, this is asking for too much from those who merit more, while if it endorses only the principle of responsibility, it is asking for too much from those who are in need.

To say that group A is asking for too much from group B

¹²⁹ – An agreement that reflects the criterion of impartiality.

implies that the interests of group B are not given equal treatment, hence asking for too much is an infringement of the conditions of impartiality. In order to embrace the requirement of impartiality, a theory of justice based on the idea of a reasonable agreement must necessarily aspire to reconcile the principles of compensation and responsibility. In this chapter, I have attempted to show how this reconciliation may be achieved.

7. ON LIBERAL NEUTRALITY

Contemporary liberalism has long been fascinated with the idea of neutrality. Many contemporary advocates of liberalism assume that the neutrality of the liberal state respect individual conceptions of the good is one element distinguishing liberalism from all other political traditions¹³⁰. Yet lately things have started to changed, and doubt has finally penetrated this seemingly indisputable liberal postulate. Indeed over the last few years a growing number of scholars, both liberal and non-liberal, have started questioning the validity of this assumption. Is neutrality a red herring in the lexicon of liberal political theory? Should liberalism abandon any claim for neutrality? This chapter will consider these questions.

In 7.I, by exploring the close affinity between neutrality and impartiality, I will argue that neutrality is indeed one of the fundamental pillars of the liberal project. Yet it is also important to emphasize that the concept of neutrality is only one

¹³⁰ – Although the term 'neutrality' is a fairly recent addition to the lexicon of liberal concepts, I believe the suggestion by Goodin & Reeve that "similar notions have long figured largely in recognizably liberal political discourses" is fundamentally correct; see Goodin and Reeve (1989), p.1.

of the key liberal concept, therefore it is wrong to assume that the condition of neutrality is sufficient to define liberalism.

In 7.II, I will consider some of the criticisms that have been made to the concept of liberal neutrality. In particular I will argue that commentators have wrongly assumed that neutrality is consequential neutrality, and therefore impossible to put in practice. I believe that liberals sympathetic to neutrality do not endorse consequential neutrality, hence the charge that neutrality is quixotic or illogical is misplaced.

In 7.III, I will put forward a conception of neutrality based on the contractualist theory of reasonable agreement. I will argue that neutrality refers to a system of rules which reflects as a whole a set of values that everyone could accept. This account of neutrality is contractarian to the extent that such system of rules must be compatible with the idea of an agreement that no one could reasonably reject.

In 7.IV, I will test this contractualist account of neutrality on what is arguably the most delicate question faced by a neutral state, namely, how to deal with illiberal conceptions of the good. Do illiberal conceptions of the good deserve the same treatment as liberal conceptions of the good? I will argue that within the boundaries of a neutral state, illiberal conceptions of the good can be dealt with adequately by the contractualist theory of neutrality as reasonable agreement.

In 7.V, I will switch the above question concerning illiberal conceptions of the good to the international context. That is to

say, how should a liberal state committed to neutrality react when it is faced by threats from illiberal political cultures? I will argue that in these cases, the liberal state is best advised to leave neutrality behind and dogmatically uphold a liberal conception of the good. Part 7.VI will provide a brief summary of the main arguments put forward in this chapter.

7.I. Neutrality and Impartiality.

In Patterns of Moral Complexity Larmore defines the concept of neutrality in the following terms:

The state should not seek to promote any particular conception of the good because of its presumed *intrinsic* superiority – that is, because it is supposedly a *truer* conception. (Larmore 1987, p.43; emphasis in original)

I suggest we adopt this definition as the primary sense of neutrality, partly because this definition has proved to be rather successful, being adopted by a number of scholars over the last few years¹³¹, and partly because other definitions of neutrality, as we shall see, portray a conception of liberal neutrality that no liberal would in fact endorse.

It is interesting to note that the above conception of

¹³¹ – See Rawls (1993a), pp.190-195; Lehning (1991).

neutrality has strong similarities with the concept of impartiality. Throughout this thesis I have assumed a specific definition of impartiality, where impartiality refers to the belief that the well-being of all moral beings matters intrinsically, therefore to be impartial implies giving all interests equal weight, or in other words impartiality is about taking equal account of the interests of all the parties.

It is rather obvious that this condition of impartiality would not be respected if the state was not neutral. In other words if a state promotes a particular conception of the good because of its presumed *intrinsic* superiority, then the conditions of impartiality would be infringed, since the interests of those with different conceptions of the good would not be given equal weight. I believe that it is precisely because of the close affinity between neutrality and impartiality that the former concept is an important qualifying assumption of liberalism in general, and of the liberal state in particular.

Although the concepts of neutrality and impartiality are closely associated (in common usage these two terms are often used synonymously), I believe that it is possible to draw a distinction between these two terms. Indeed, for reasons we shall see later, a distinction between neutrality and impartiality at the analytical level is critically important. The fundamental difference between impartiality and neutrality is the following: while neutrality applies to the sphere of politics, in particular to political procedures, impartiality refers to the moral sphere, to the Kantian

claim of inter-subjectivity. Impartiality is, in a sense, the essence of the categorical imperative.

In political theory, the concept of neutrality is something of a term-of-art: it is a political term that applies exclusively to the state, while impartiality applies to the more general idea of non-favouritism. Being concerned with the impact of political institutions on individual citizens, liberals are only interested in the way neutrality applies to the sphere of the state. For a liberal neutrality is political neutrality; neutrality concerns the procedures that establish the relationship between the state and its citizens. As Larmore points out:

neutrality as a political ideal governs the *public* relations between persons and the state, and not the *private* relations between persons and other institutions. (Larmore 1987, p.45; emphasis in original)

For a liberal, neutrality is a political, not a general social ideal. (Larmore 1987, p.46)

The political characteristic of neutrality contradicts the views of some commentators who wrongly assume that neutrality applies to any circumstance "in so far as one is in a position to exercise some form of influence" (Montefiore 1975, p.5), whether it applies to the state and the citizen, or to parents and their

children¹³².

If neutrality is a political concept, impartiality is a moral motivation. Indeed this is the way the concept of impartiality has been used recently by advocates of justice as impartiality, from Scanlon who defines impartiality in terms of "the desire to be able to justify one's actions to others on grounds they could not reasonably reject" (Scanlon 1982, p.116), to Kymlicka's idea that "all human beings have a good, and their well-being matters intrinsically" (Kymlicka 1990b, p.111), to Barry's idea that impartiality implies "tak[ing] an equal account of the interests of all the parties" (Barry 1989b, p.269).

The reason why it is important to distinguish the concept of neutrality from that of impartiality is because it is not always practical to endorse neutrality: for example when a liberal state is threatened by other illiberal states it is better to give up on neutrality altogether¹³³. Yet neglecting neutrality does not mean that a liberal must give up on impartiality, since impartiality and neutrality appertain to different spheres. Indeed in some extreme cases it is necessary to sacrifice neutrality in order to preserve impartiality.

The place of neutrality in the liberal tradition has come under serious scrutiny in the last few years, and although Rawls (1988 & 1989) and Ackerman (1980) argue that neutrality is at the

¹³² – Raz adopts Montefiore's definition of neutrality, as well as the example of a father intervening in some dispute between two children. See Montefiore (1975, p.7) and Raz (1986), p.114.

¹³³ – I will deal with this question in 7.V below.

heart of contemporary liberalism, other liberal theorists deny neutrality any place in the liberal tradition, although they claim to be liberals; in this fast growing family we find Alexander & Schwarzschild (1987), Bellamy (1989) and Raz (1986).

The critiques of neutrality stemming from sympathizers of the liberal project all share a common characteristic, namely, they think that neutrality is consequential and therefore illogical. That is to say, the concept of neutrality is rejected to the extent that it contradicts other fundamental liberal assumptions. In what follows, I want to reject this critique of neutrality.

7.II. The Impossibility of Neutrality.

The impossibility of neutrality refers to the charge that the concept of liberal neutrality is infested by internal contradictions. These contradictions can take different forms: the universalistic and objective nature of neutrality are either seen as detrimental to liberalism, or alternatively it is claimed that the ends of neutrality cannot be put into practice.

I suggest we start by looking at the tendency of equating neutrality with universalistic claims. Both Alexander & Schwarzschild, and Bellamy, seem to equate the concept of neutrality with universalistic claims reflecting an objective nature; neutrality is universal since it is believed to take everyone's position into account, furthermore it is seen as

objective to the extent that it is not contaminated by any kind of preferences.

To the extent that liberalism rests on an abstract principle of neutrality, Alexander & Schwarzschild claim that neutrality is detrimental to liberalism. In order to expose the inherent contradictions of neutrality, and its detrimental influence on liberalism, Alexander & Schwarzschild undertake a critical analysis of Dworkin's neutrality-based liberalism.

In striving to distinguish expensive tastes from basic needs, Dworkin is charged with attempting to overcome a fundamental contradiction inherent in the concept of neutrality:

Dworkin cannot rely on such distinctions as that between 'objective needs' and 'subjective wants'; or that between 'normal' physical limitations and 'abnormal' physical disabilities. These distinctions rely on value judgments; hence Dworkin cannot rely on them without violating his own constraint of neutrality. (Alexander & Schwarzschild 1987, p.101)

For anyone who takes neutrality seriously, handicaps, cravings, and expensive tastes are on a par with one another. (Alexander & Schwarzschild 1987, p.101)

Following the logic of their argument, Alexander & Schwarzschild are led to the conclusion that there is no neutral

way of measuring resource holdings, hence liberalism is advised to renounce to neutrality for its theoretical grounding.

In an argument reminiscent of Alexander & Schwarzschild, Bellamy claims that neutrality is characterised by an intrinsic paradox, since it is impossible to separate the right from the good. As in the case of Alexander & Schwarzschild, Dworkin's neutrality is the target of Bellamy's critique, indeed Bellamy detects a contradiction between Dworkin's claim for neutrality and the necessity for liberalism to exclude some conceptions of the good:

Once Dworkin acknowledges that non-neutral outcomes can have repugnant consequences, how can he devise methods to avoid only some of them without himself infringing neutrality[?].
(Bellamy 1989, p.29)

It cannot be denied that there are some problems with Dworkin's account of neutrality¹³⁴. Yet all this tells us is that there is a problem with Dworkin's account of neutrality and liberalism, rather than with the concept of neutrality *per se*. In other words by criticising Dworkin's account of neutrality Alexander & Schwarzschild and Bellamy are not defeating the concept of neutrality, instead they are only defeating Dworkin's definition of neutrality.

¹³⁴ – In his recent Tanner lecture Dworkin recognises that there are some serious flaws with his previous account of liberalism and neutrality. See Dworkin (1990) p.10n.

Neutrality does not mean that no form of inequality can ever be justified. If we adopt this conception of neutrality, then the state should for example distribute wheel chairs to anyone who asks for one, whether they are paraplegic or simply lazy people who want to move from their TV couch to the kitchen without getting up on their feet. Intuitively this cannot be what neutrality is about. To most liberals neutrality means that all conceptions of the good should be treated according to the same rules, and that decisions concerning similar demands must follow similar procedures, although these procedures can lead to egalitarian results. In other words neutrality refers to the decision procedures, and not the outcomes.

If we return for a moment on Bellamy's critique of Dworkin, it is interesting to point out that Bellamy assumes that neutrality ought to be judged from a consequentialist perspective. Indeed most criticisms of neutrality are based on the assumption that neutrality refers to the consequences of actions. Clearly this is the view of Raz, who is happy to endorse Montefiore's claim that "To be neutral in any conflict is to do one's best to help or hinder the various parties concerned in an equal degree" (Montefiore 1975, p.5). Raz's consequentialist view of neutrality has since been adopted by other liberal perfectionists, who argue that liberalism ought to abandon neutrality and embrace perfectionism¹³⁵. In what follows I want to consider and reject Raz's views on consequentialist neutrality.

According to Raz the problem with neutrality lies in the idea

¹³⁵ – See Caney (1991).

of allocating influence "in an equal degree", in fact he argues that providing help or hindrance *in any degree* will never secure that the outcome will influence the receiving social actors to an equal degree. Furthermore Raz argues that even by abstaining to help, a neutral party may in fact hinder the weaker of the conflicting parties; in other words if we assume an unequal balance of powers, not altering the status quo is like favouring the stronger party. The conclusion Raz draws from all this is that we are better off abandoning the dream of neutrality since its theoretical appeal cannot be translated to practice.

While Raz's critique of Montefiore's conception of neutrality cannot be refuted, it would be erroneous to assume that Raz's analysis constitutes the definitive critique of neutrality; paradoxically, I feel the concept of neutrality comes out stronger than it was before it fell under the analytical microscope of Raz. In fact contrary to what Raz seems to think, his critique of neutrality does not undermine the validity of the concept *per se*, but at most it undermines the validity of Montefiore's principle of neutrality.

By adopting Montefiore's dictum on neutrality as a primary sense of neutrality, Raz manages to discredit what is a weak and unpersuasive account of neutrality: consequentialist neutrality. It follows that to the extent that Raz has practically obliterated Montefiore's account, the concept of neutrality is in many ways stronger now that it is finally free of Montefiore's heavy mantle. In other words there is no reason why we have to stick with

Montefiore's dictum as the primary sense of neutrality. Montefiore's dictum is consequentialist, yet there is more to neutrality than consequentialist neutrality.

As Kymlicka points out, Raz is simply wrong to assume that liberals who defend neutrality endorse consequentialist neutrality. Instead according to Kymlicka, liberals like Rawls endorse justificatory neutrality, or the view that "the state does not take a stand on which ways of life are most worth living, and the desire to help one way of life over another is precluded as a justification of government action" (Kymlicka 1989b, pp.883-884).

In his most extensive exposition on the subject of neutrality, John Rawls (1988) spells out the difference between neutrality of aim and neutrality of effect or influence, and clearly takes side with the former view. Neutrality of aim claims that "that the state is not to do anything intended to favour or promote any particular comprehensive doctrine rather than another, or to give greater assistance to those who pursue it" (Rawls 1988, p.262), while neutrality of effect or influence claims that "the state is not to do anything that makes it more likely that individuals will accept any particular conception rather than another unless steps are taken to cancel, or to compensate for, the effects of policies that do this" (Rawls 1988, p.262).

As Rawls rightly points out, political liberalism abandons neutrality of effect or influence for being 'impracticable', endorsing in its place neutrality of aim. This is important because it shows that Rawls would have no problem accepting the critiques

aimed at the concept of neutrality by Alexander & Schwarzschild, Bellamy or Raz, since Rawls is the first to admit that political liberalism cannot treat all conceptions of the good equally (what Montefiore referred to as helping or hindering "in an equal degree"), hence neutrality of effect or influence is both impracticable and undesirable.

So far I have argued that neutrality cannot be rejected for being illogical, since neutrality does not imply consequential neutrality. In what follows, I want to give a different account of neutrality, namely an account which is based on the contractualist theory of reasonable agreement.

7.III. Neutrality and Reasonable Agreement.

Liberal neutrality is not about consequences. When a liberal invokes the neutrality of the state, equality of effects or influence is not what is being defended. Instead liberals endorse justificatory neutrality rather than consequential neutrality. But what exactly does justificatory neutrality mean?

It seems to me to be irrefutable that justificatory neutrality refers in the first place to the procedures that establish how conflicts between conceptions of the good are to be resolved¹³⁶.

¹³⁶ - Although proceduralism is part of neutrality, later in this chapter I will argue that in fact proceduralism has a dark side, indeed as we shall see it is best to divorce the concept of neutrality from all references to proceduralism.

Thus neutrality is about equality of treatment. As Jones points out:

A neutral state is one that deals impartially with its citizens and which remains neutral on the issue of what sort of lives they should lead. Those who endorse the idea of the neutral state hold that it is not the function of the state to impose the pursuit of any particular set of ends upon its citizens. Rather the state should leave its citizens to set their own goals, to shape their own lives, and should confine itself to establishing arrangements which allow each citizen to pursue his own goals as he sees fit - consistent with every other citizen's being able to do the same. (Jones 1989b, p.9; emphasis added)

It seems to me that the key aspect of Jones's account of neutrality is the emphasis on proceduralism. Jones refers to such procedures in terms of "establishing arrangements" which allow everyone to pursue their conception of the good on the condition that others are not harmed in the process. The procedural dimension of neutrality has also been emphasized by others, in particular by Larmore who claims that:

[the neutrality of the liberal state] is not meant to be one of outcome, but rather one of procedure. That is, political neutrality consists in a constraint on what factors can be

invoked to justify a political decision. Such a decision can count as neutral only if it can be justified without appealing to the presumed intrinsic superiority of any particular conception of the good life. (Larmore 1987, p.44; emphasis in original)

There are two reasons why proceduralism has been emphasized as a central aspect of neutrality. First of all, because fair proceduralism reflects the conditions of impartiality. Although earlier on I argued for a conceptual separation between the concepts of neutrality and impartiality, there is an important sense in which treating everyone as equals is confirmation that the interests of all are given equal account, hence that the well-being of all human beings matters intrinsically. Although Jones does not tell us what he means by impartiality, his claim that "a neutral state is one that deals impartially with its citizens" is essentially correct.

Secondly, by equating neutrality with fair procedures, the limits of the claim that neutrality is the core concept in the liberal tradition become apparent. In other words, if neutrality is a characteristic of liberal procedures, then neutrality cannot be the only or main foundation of the liberal project: just as proceduralism is a necessary but insufficient condition for social justice, neutrality is a necessary but insufficient condition for liberalism.

Although proceduralism is important for an account of

neutrality, the idea of proceduralism is extremely vague and it can easily lead to a misinterpretation of neutrality. For example Peter de Marneffe points out that neutrality can be understood in two different ways: as constitutional neutrality or as legislative neutrality. Constitutional neutrality means that "a system of laws is neutral if, as a whole, it can be justified solely in terms of neutral values", while legislative neutrality means that "a system of laws is neutral when there is no law which cannot itself be justified in terms of neutral values (or: for every law, there is a neutral reason which warrants it)"¹³⁷.

Both constitutional and legislative neutrality endorse a form of proceduralism, although most liberal sympathizers of neutrality would only endorse constitutional neutrality¹³⁸. The problem with legislative neutrality is that it does not guarantee that the conditions of impartiality will be respected, in fact if every law is justified in terms of neutral values, it is very likely that the interests of some groups will suffer. The distinction between constitutional and legislative neutrality is important because, as de Marneffe points out "violations of legislative neutrality may be fully compatible with constitutional neutrality" (de Marneffe 1990, p.259).

Furthermore, pure procedural notions are insensitive to the

¹³⁷ - These definitions of constitutional and legislative neutrality are taken from de Marneffe's unpublished doctoral thesis Liberalism and Education, quoted in Lehning (1991). For a similar account of constitutional and legislative neutrality, see de Marneffe (1990), p.259.

¹³⁸ - With the possible exception of Dworkin.

distinction between personal and external preferences. If by pure proceduralism we understand a mechanism whereby the preferences of everyone are treated with equal concern and respect, then neutrality adds up to nothing more than legislative neutrality. The problem with legislative neutrality is that it precludes the possibility to make a principled or substantive judgement, which is why constitutional neutrality is preferred¹³⁹.

The ambiguity with the concept of proceduralism probably explains why Rawls made some subtle but important changes in rewriting his 1988 article "The Priority of Right and Ideas of the Good" as Lecture V of his Political Liberalism. In fact we find that in the latter exposition all references to proceduralism are eliminated. Compare the following passages:

As a political conception for the basic structure justice as fairness as a whole can be seen as exemplifying a kind of procedural neutrality, and it also hopes to satisfy neutrality of aim in the sense that the basic institutions and public policy are not to be designed to favor any particular comprehensive doctrine. (Rawls 1988, p.263)

As a political conception for the basic structure justice as

¹³⁹ - On the question of pure proceduralism and the personal/external preference approach, see de Marneffe (1990), pp.262-264. I will return to the principled or substantive approach to neutrality in part 7.VI below.

fairness as a whole tries to provide common ground as the focus of an overlapping consensus. It also hopes to satisfy neutrality of aim in the sense that basic institutions and public policy are not to be designed to favor any particular comprehensive doctrine. (Rawls 1993a, pp.193-194).

Clearly in 1993 Rawls excludes all references between justice as fairness and procedural neutrality which he made in 1988. Another indicative example is the following: in 1988 Rawls writes that "Justice as fairness is not, without important qualifications, procedurally neutral" (Rawls 1988, p.261), although in 1993 he simply writes that "Justice as fairness is not procedurally neutral" (Rawls 1993a, p.192).

The point Rawls wants to make in both the 1988 article and in Political Liberalism is that the principles of justice as fairness are substantive and express far more than procedural values, and so do its political conceptions of person and society (Rawls 1988, p.261; 1993a, p.192)¹⁴⁰. It seems to me that de Marneffe's distinction between constitutional and legislative neutrality captures in important ways Rawls's distinction between the substantive element of justice as fairness and simple proceduralism, hence in what follows I suggest we consider neutrality in terms of constitutional neutrality.

I believe that by reading neutrality in terms of

¹⁴⁰ - In Political Liberalism Rawls adds that the political conceptions of society and persons "are represented in the original position" (Rawls 1993a, p.192).

constitutional neutrality enables us to capture the way in which the account of neutrality can be based on the contractualist theory of reasonable agreement. We saw earlier that according to Jones, the state should confine itself to establishing arrangements which allow each citizen to pursue his or her own goals as he or she sees fit, yet Jones fails to tell us how such arrangements are to be established. The answer lies in the contractualist theory of reasonable agreement. Indeed I believe we can go as far as saying that de Marneffe's definition of constitutional neutrality can be rephrased in the following terms: a system of laws is neutral if, as a whole, it can be justified in terms of a reasonable agreement (an agreement on terms that no one could reasonably reject).

In Chapter 3 we saw that according to Scanlon, contractualism is about establishing a system of rules for the general regulation of behaviour which no one could reasonably reject as a basis for informed, unforced general agreement¹⁴¹. It follows that neutrality refers to the procedure or set of rules that is based on values that no-one could reject. In order to establish the set of rules based on values that no-one could reject, I believe it is necessary to refer to the two key notions of reasonable agreement: the idea of agreement and the concept of reasonableness as 'not asking for too much'.

Thus according to the contractualist theory of reasonable agreement, the system of rules which determines the neutrality of the state must not reflect the condition of bargaining advantages

¹⁴¹ – See Scanlon (1982), p.110.

or mutual advantage, instead the system of rules must be one that is accepted by those who are moved by the *idea of agreement*, or in other words by those who are prepared to revise their conception of the good in order to make room for others.

Furthermore, the role of reasonableness, defined by the maxim 'not asking for too much', acts as the rationale that illuminates our moral intuitions. Thus the system of rules is determined by an agreement characterised by the fact that the parties in pursuing their conceptions of the good will not be asking for too much.

To recapitulate the argument so far, the account of neutrality given by the contractualist theory of reasonable agreement focuses on the procedures established by a set of rules based on values that no one could reject, where this set of rules depends upon a prior notion of justice.

In order to test the validity of neutrality based on the contractualist theory of reasonable agreement, I suggest we consider how this account of neutrality deals with arguably the most difficult question for any theory of neutrality, namely, how it deals with illiberal conceptions of the good. It is to this question that I want to move to next.

7.IV. Neutrality and Illiberal Conceptions of the Good.

In what follows, I will consider how the account of neutrality based on the contractualist theory of reasonable agreement responds

to the well known case of the American Nazis staging a demonstration in the town of Skokie, Illinois. In January 1978, the Illinois Supreme Court (in a 7 to 1 decision) ruled in favour of Frank Collins, the leader of the National Socialist Party of America (NSPA), who had announced that his Nazi party was going to stage a demonstration in the town of Skokie, Illinois. Skokie is characterised by that fact that it has the highest number of holocaust survivors of any city in the United States outside of New York city.

The case of the NSPA and Skokie is particularly interesting because it raises the question of whether a liberal, neutral state ought to defend the rights of freedom of speech of the NSPA, or whether in this case liberalism can justify the curbing of free speech without contradicting its neutrality. What is interesting to note is that the main argument behind the Illinois Supreme Court decision was the so-called 'content neutrality rule' according to which political speech shall not be abridged because of its content. Furthermore at the time the American Civil Rights Union felt that adherents of the NSPA had rights as any other American citizen, hence liberalism had a duty to defend their right.

In a recent article, Raphael Cohen-Almagor (1993) argues that the decision of the Illinois Supreme Court was wrong, indeed that curbing the rights of the NSPA in that occasion can be justified in liberal terms. In what follows I will argue that although Cohen-Algamor's conclusion is generally correct, his reasoning is not totally convincing. In fact I believe that applying the analysis of

liberal neutrality and reasonable agreement to the case of Skokie leads to a different appraisal.

According to Cohen-Almagor, the constraints on the rights of speech can be justified on two accounts: by the Harm Principle and the Offence Principle. Under the Harm Principle, restrictions on liberty may be prescribed when there are sheer threats of immediate violence against some individuals or groups. Under the Offence Principle, psychological offence is morally on a par with physical harm, hence any form of expression that is designed to inflict psychological harm can be curbed. In the case of Skokie, Cohen-Almagor argues that the Offence Principle can be invoked to constraint the freedom of speech of the NSPA.

Of course Cohen-Almagor does not say that offence is sufficient to curb freedom of speech, in fact applying this principle rigidly could back-fire. For example, Southern whites could ban civil rights marches, especially those that are held by blacks, on the grounds that they find the offensive:

The fact that some individuals are offended by a speech which advocates equal rights is not sufficient reason for its restriction. The [Offence] Principle affects freedom of expression when the speech in question contradicts fundamental background rights to human dignity and equality of concern and respect. (Cohen-Almagor 1993, p.469-470)

While I agree with the general line of Cohen-Almagor's

argument, I find some of his supporting claims confusing. For example, according to Cohen-Almagor the case of Skokie is particularly interesting because the Jews living in the area did not have a choice to avoid the offense, hence the choice of location undermines the legitimacy of the march. In other words Cohen-Almagor claims that if it was organized in any other part of the country, on ethical grounds the march could not (and should not I presume) have been prevented from taking place. But choosing Skokie as the location of the march underlines specific intentions and motives, hence according to Cohen-Almagor we can appeal to the Offence Principle to curb the NSPA's freedom of speech.

It seems to me intuitively wrong to make a moral argument against a Nazi group revolve in the last analysis around a geographical question. Ethics does not have geographical boundaries, hence the fact that the NSPA's march was organized in Skokie should not make any substantive differences.

By grounding his argument on the issue of the geographical location of the march, Cohen-Almagor is contradicting his previous claim that the speech in question must not refute fundamental background rights to human dignity and equality of concern and respect. Such background rights are refuted whether the march is held in Skokie (Illinois) or Tuscaloosa (Alabama)¹⁴². It is not the location that determines whether the march is offensive or not, instead it is the very content of the Nazi conception of the good

¹⁴² - I don't know much about this town, but I assume that it does not have a concentration of Jews or survivors of the holocaust above the average of any other American city.

that is problematic. The question we must ask ourselves is whether banning Nazi propaganda or preventing the NSPA from staging their demonstration is compatible with the endorsement of neutrality. I believe it is, for the following reason.

Clearly, if the Nazi demonstration is banned on the assumption that it is an inferior conception of the good, or alternatively because the state favour a non-Nazi conception of the good that it believes to be intrinsically superior, than the very essence of neutrality would be infringed. Yet endorsing neutrality does not mean that Nazi propaganda cannot be stopped. I argued earlier that neutrality refers to a system of rules that no one could reject, where this system of rules depends upon a prior notion of justice. It follows that under neutrality, Nazi propaganda would be banned by a set of rules that no one could reject as a basis for informed, unforced general agreement.

The contractualist account of neutrality indicates that the problem with the decision of the Illinois Supreme Court lies with the 'content neutrality rule'. To use de Marneffe's terminology, it would appear that the Illinois Supreme Court endorsed legislative neutrality, in fact the 'content neutrality rule' (that political speech shall not be abridged because of its content) reflects the idea behind legislative neutrality that every rule, in this case the rule concerning freedom of speech, must itself be justified in terms of neutral values.

By following a reasoning similar to that of the Illinois Supreme Court, the ACRU (American Civil Rights Union) was culpable

of making the same confusion between constitutional and legislative neutrality. According to the ACRU, everyone has a right to equal treatment, even a Nazi, hence a liberal state cannot prevent the Nazi group from enjoying the same rights as any other American citizen.

The type of reasoning exemplified by the ACRU's reflects a concern for legislative neutrality (what I referred to earlier as simple or pure proceduralism) that is not endorsed by the contractualist account of neutrality based on reasonable agreement. According to the latter, if the system of rules reflects the condition of reasonable agreement, then forms of illiberal propaganda can be repressed. The reason why the system of rules as a whole would curb the rights of the NSPA is because, as Waldron (1989) points out, in their content and tendency the Nazi's speeches are incompatible with the very idea of the right to freedom of expression they are asserting¹⁴³.

7.V. Neutrality and Illiberal States.

So far I have considered neutrality as applying to the internal arrangements of the state, or in other words I have assumed that neutrality concerns the political relationship between the state and the citizen. Yet there is another instance in which

¹⁴³ - For an illuminating treatment of the argument that Nazis do not have a right to freedom of expression, see Jeremy Waldron (1989).

the phrase 'neutral state' applies, namely, in the relationship between states. Thus for example, when in 1992 Yeltsin proclaimed the Commonwealth of Independent States, dismissing the Soviet Union and Gorbachov to the dustbin of history, President Bush declared that the United States of America would take a 'neutral' stance on this matter.

Some commentators have used analogies concerning the international context in order to highlight what they see as neutrality's contradictory nature. Thus one of the examples Raz utilizes is that of a country (say White) maintaining commercial ties with another country (Red) who is engaged in a war with yet another country (Blue). Raz uses this example to show the impossibility of applying the principle of neutrality, since according to Raz failure to help is hindering, and it is impossible to help or hinder to an equal degree.

It is indisputable that if, as a result of the war between Reds and Blues, Whites refuses to maintain commercial ties with Reds (since Whites clearly favors Blue), then Whites is violating the idea of neutrality. Yet it seems to me that the interesting question is not whether Whites can help or hinder the other countries to an equal degree¹⁴⁴, but if Whites has a moral obligation to endorse neutrality in the international context. If for example we assume that Blues are under the leadership of a dictator who is waging a war against the Reds as a first step

¹⁴⁴ - I have already argued that this account is based on Montefiore's erroneous definition of neutrality.

towards world domination, should Whites adopt a position of neutrality? In other words, how should a liberal state respond to the threat of non-liberal states? Does neutrality between states follow the same reasoning as between the liberal state and conceptions of the good of its citizens?

In what follows, I want to set the foundations for the following argument: questions of neutrality of the state only apply to internal considerations, while in the case of international relations, the liberal state is best advised to drop neutrality, and to defend liberal values as dogmatically as non-liberals defend their positions. Needless to say that a detailed and satisfying defence of the above claim would deserve a thesis of its own, indeed I do not even claim to have the expertise or competence to deal adequately with the ethical issues of international relations. My intention here is simply to lay out possible avenues that may be followed if the argument I have made thus far concerning contractualism and neutrality is to be applied to the international context. In particular, I want to emphasize the following two considerations.

First, taking a neutral stance on all international issues may contradict one of the fundamental pillars of liberal state, namely, providing security to its citizens. If the liberal state refuses to intervene against illiberal states (for example against Serbian atrocities in their conflict against the Muslims, or against Algeria's Muslim fundamentalists), liberal nations may find themselves outnumbered by their foes, with the risk of endangering

its own citizens.

Secondly, it seems to me that the very conditions for adopting a neutralist stand are missing from the international context, hence it would be foolish for a liberal state to take a neutral stance on principle. As I said before, by neutrality we must understand constitutional neutrality, whereby a system of laws as a whole can be justified in terms of neutral values. Yet for the moment the system of rules governing the relations between states is not adequate, especially as international rules can be broken with little fear of retribution, hence from an ethical perspective neutrality cannot apply. As a result, there is no reason for a liberal state to feel under an obligation to endorse neutrality.

To summarize, it seems to me that when confronted with other hierarchical societies, the liberal state is best advised to leave neutrality at home. As Barry (1990b) rightly points out, belief in neutrality is not going to cut much ice with a non-liberal, since a non-liberal "is not simply someone who adheres to a dogma, but someone who adheres to it dogmatically" (Barry 1990b, p.13). It follows that in the international arena the liberal state should defend its liberal values dogmatically, while reserving neutrality for internal disputes.

7.VI. Conclusion.

In the remaining pages of this chapter, I want to summarize

the central concepts of the contractualist account of neutrality I have defended. First of all, the contractualist account of neutrality expresses far more than procedural values, in fact like in case of Rawls, contractualism gives a substantive account of neutrality.

Although Rawls's liberalism embraces neutrality, he does not want to treat all conceptions of the good equally, in fact he points out that political liberalism:

must impose restrictions on permissible comprehensive views, and the basic institutions ... [must] inevitably encourage some ways of life and discourage others, or even exclude them altogether. (Rawls 1988, p.264)

The core intuition behind Rawls's idea is that neutrality "sets limits to permissible ways of life" (Rawls 1988, p.251). According to Rawls, neutrality performs the crucial function of "impos[ing] restrictions on permissible comprehensive views" (Rawls 1988, p.264), which implies that "the basic institutions ... inevitably encourage some ways of life and discourage others, or even exclude them altogether" (Rawls 1988, p.264).

In order to distinguish between permissible and non-permissible conceptions of the good, it is necessary to distinguish neutrality from pure proceduralism, and search for the moral grounds of neutrality in the principles of justice that regulate the basic social and political institutions. Thus echoing Rawls,

according to de Marneffe neutrality is when:

the principles of justice that regulate basic social and political institutions [are] justifiable in terms of moral and political values that any reasonable person would accept as the basis of moral claims regardless of his or her particular conception of the good. (de Marneffe 1990, p.253)¹⁴⁵

The contractualist theory of neutrality follows the same approach as Rawls and de Marneffe. The difference is that in the contractualist theory of reasonable agreement, the substantive element (moral and political values) in the account of neutrality are explicitly determined by the idea of agreement and the condition of reasonableness. That is to say, the contractualist framework of a reasonable agreement is needed in order to formulate the values that everyone would find acceptable.

Although the concept of neutrality is, in important ways, related to the principles of justice, neutrality has also important links with the ideal of political legitimacy. That there is an affinity between the concepts of neutrality and legitimacy is a view endorsed by Kymlicka and de Marneffe. Although Kymlicka explicitly writes that "there is no inherent connection between

¹⁴⁵ - de Marneffe calls this *neutrality of grounds*, which is different from *concrete neutrality*, or the principle that the state may not limit individual liberty in ways that advance one particular conception of the good.

neutrality and state legitimacy" (Kymlicka 1990a, p.237n.13), he nevertheless feels that neutrality plays a role in a theory of political legitimacy, in fact he claims that: "I believe that liberal neutrality is the most likely principle to secure public assent in societies like ours, which are diverse and historically exclusionary" (Kymlicka 1990a, p.229).

Similarly de Marneffe argues that the significance for liberalism of neutrality is best understood in light of the liberal ideal of political legitimacy, in fact "the liberal ideal of legitimacy requires that principles of justice be justifiable in terms of neutral values" (de Marneffe 1990, p.256).

I believe it is of fundamental importance to understand the concept of neutrality as acting between the principles of social justice and the idea of political legitimacy; neutrality is the go-between connecting the legitimacy of the state with principles of social justice. The reason why this is an important question is because it shows that the function of neutrality in the liberal project is not to bring about better consequences, instead neutrality has the function of giving liberalism the coherence and solidity it may appear to lack compared to fundamentalist or dogmatic moral and political theories.

I believe that a full understanding of neutrality must take account of the function of neutrality in the liberal project. In what follows I will argue that the function of neutrality in the liberal project is not primarily that of bringing about better consequences, as perfectionist theorists as Raz have argued.

Instead the function of neutrality is that of bridging the gap between the concepts of social justice and legitimacy of the state.

Defined in its most general sense, legitimacy concerns the relationship between state and subject, a relationship characterized by a belief on the part of both the state (i.e. participants in the institutions of government) and the subjects (i.e. individual citizen) that the political regime is justified and justifiable¹⁴⁶. This relationship between state and subject is grounded on consent, in fact it is generally accepted that consent is the source of political legitimacy¹⁴⁷.

What determines the consent of the state by the citizens, or in other words what makes a political regime justifiable, is ultimately settled by the theory of justice the state subscribes to. As Rawls argues in his more recent articles, a theory of justice can act as the foundation for a theory of political legitimacy:

we must stay within the limits of justice as fairness as a political conception of justice that can serve as the focus of an overlapping consensus. (Rawls 1988, p.258)

While there is a strong bind between the legitimacy of the

¹⁴⁶ - See Barker (1990).

¹⁴⁷ - As Waldron claims: "The thesis that I want to say is fundamentally liberal in this: a social and political order is illegitimate unless it is rooted in the consent of all those who have to live under it" Waldron (1987), p.140.

state and conceptions of social justice, there is a need for a filter between these two key concepts. In other words, it is necessary for conceptions of social justice to be translated into terms that form the basis of political legitimacy.

In order to bridge the gap between social justice and political legitimacy, it is important to narrow the range over which consent is sought. As Waldron (1987) rightly argues, there are two ways of interpreting the idea of an hypothetical consent. Either we look for some underlying interests and beliefs shared in common¹⁴⁸ (as in the case of Rawls's overlapping consensus), or we concede that the liberal state and its neutrality will only be appealing to those who are moved by liberal intuitions. Waldron rejects the former idea in favour of the latter.

It follows that before establishing whether state institutions have the consent of the people, we must first restrict the arena of actors whose consent the state institutions must seek: in the last analysis, only those who are moved by liberal intuitions are invited to legitimize the liberal state. In other words, those who by giving their consent legitimise the liberal state, must espouse conceptions of the good worthy of consideration for claims of social justice.

It is here that the function of neutrality, acting as the intermediate stage between legitimacy and social justice, takes full shape. Neutrality weeds out all those conceptions of the good

¹⁴⁸ - "a conception of the good life that all people, even those with the most diverse commitments, can be said to be engaged in". Waldron (1987), p.145.

that should not figure in liberal considerations of legitimacy or social justice. In fulfilling this function the concept of neutrality acts as the filter between social justice and legitimacy, as shown in Figure 7.1 below.

political legitimacy <-----> neutrality <-----> social justice

Fig. 7.1

If we accept the view that the function of neutrality is that of bridging the gap between the concepts of social justice and legitimacy of the state, and not that of bringing about better consequences, then the inadequacy of Raz's critique of neutrality becomes clear to see.

8. CONCLUSION

8.1. A Short Summary.

Twenty years after the resuscitation of the idea of contractualism at the hands of John Rawls, contractualism seems once again to be out of favour with liberal political philosophers. It cannot be denied that it is partly Rawls's own fault if contemporary liberalism is ambiguous about contractualism. After all, Rawls's attempt to justify a theory of justice on contractualist foundations proved to be more attractive in theory than it was in its application.

In *A Theory of Justice* Rawls re-invents the contractualist approach in terms of the original position, a hypothetical condition that Rawls compares to the state of nature in classical accounts of the social contract. Unfortunately, Rawls's use of contractualism in general and his account of original position in particular have proved to be the most discussed, arguably the most fascinating, but surely the least convincing aspect of his theory of justice.

In Chapter 2 I argued that there are two key notions in the contractarian discourse, namely 'agreement' and 'cooperation'. I

also argued that Rawls's treatment of these two key notions is not convincing, in fact it appears that Rawls is appealing to two opposing and irreconcilable interpretations of the social contract: the Kantian strategy, where a morally loaded conception of agreement as impartiality is the crucial moment in a contractarian approach, and the Hobbesian strategy, where the idea of social cooperation for mutual advantage is the determining factor in the drafting of a social contract. The argument presented in this thesis revealing Rawls's two social contract theories is further confirmation of Barry's (1989b) claim that there are two theories of justice in Rawls's work.

Notwithstanding all these complications relating to Rawls's attempt to breathe new life in the neglected tradition of the social contract, I believe that contractualism is still the right answer to contemporary ethical and political problems afflicting liberalism. Indeed the aim of this thesis was to defend the role of contractualism in contemporary liberal political theory.

It seems to me that the problem with Rawls's political theory (indeed with liberal political theory in general) does not lie with contractualism, instead it is the limited use liberalism makes of contractualism that is most problematic. Rawls considered contractualism as an approach valid at the ethical level, that is, as a device that can help us to determine right actions from wrong actions. Here lies the problem with Rawls's treatment of the social contract. It seems to me that what Rawls should have done is to raise contractualism from the ethical level to the meta-ethical

level. At the meta-ethical level, contractualism can help us with our inquiries concerning the nature of morality, by facilitating the process by which we come to terms with our moral judgments and our moral intuitions.

The idea of pitching contractualism at the meta-ethical level is denoted by the theory of 'reasonable agreement'. This theory finds its first and best account in Thomas Scanlon's "Contractualism and Utilitarianism", arguably one of the most widely read and influential articles in political philosophy of the last decade. Chapter 3 comprised of a critical assessment of Scanlon's contractualism. Following in the footsteps of Scanlon, in this thesis I have attempted to consolidate Scanlon's project of raising contractualism at the meta-ethical level by emphasizing (and hopefully clarifying) two central aspects in the notion of 'reasonable agreement': the *idea of agreement* and the *concept of reasonableness*.

Endorsing the *idea of agreement* implies that a person is seeking an agreement with others on grounds they could not reject. In other words, it means that in seeking an agreement with others this person is willing to compromise or revise his or her conception of the good in order to find an agreement. The *idea of agreement* is important because it reminds us that a social contract is not exclusively or primarily about the advantageous consequences associated with reaching the agreement. Instead, what is important in a social contract is that the conditions of impartiality are respected. The *idea of seeking a reasonable agreement* is grounded

on the motivation of impartiality, since the willingness to compromise or in some way revise one's conception of the good in order to make room for others reflects our moral duty to respect the equal claim to consideration of all human beings.

In the theory of reasonable agreement, the agreement is qualified by the concept of reasonableness. It is surprising to find that although many political philosophers have appealed to the idea of a reasonable agreement in their writings, the concept of reasonableness has never been adequately discussed or defined, with the result that the concept of reasonableness is surrounded by a cloud of ambiguity, equivocation and much confusion. Starting from the assumption that there is a need to define the concept of reasonableness, in Chapter 4 I argued that the concept of reasonableness ought to be defined by the maxim 'not asking for too much'.

To recapitulate, the contractualist theory of reasonable agreement I have defended in this thesis is an attempt to combine the idea of agreement and the concept of reasonableness in order to illuminate and strengthen the moral motivation of impartiality. Of course it ought to be remembered that the intention behind this thesis is not confined to the scholastic exercise of analysing key concepts in moral and political philosophy. Instead this thesis is moved by the more general endeavour of justifying the foundations of the liberal project, and to fortify the liberal tradition from possible threats.

In order to show that the theory of reasonable agreement can

serve as a foundation for liberalism, it must be shown that the theory of reasonable agreement is equipped to respond to the challenge facing liberalism¹⁴⁹. It is to this question that I want to turn now.

8.II. Reasonable Agreement and the Challenge Facing Liberalism.

In Chapter 1 I argued that the ethical challenge facing liberalism concerns providing the conditions for the peaceful coexistence of diverse and sometimes conflicting conceptions of the good, ranging from religious beliefs to economic self-interest. It is important to emphasize that diversity and conflict are the essence of liberalism, hence the solution to the challenge facing liberalism cannot be to establish a hierarchical society¹⁵⁰ which eradicates diversity and suppresses conflict; the challenge facing liberalism can only be met by an ethical argument that reconciles diversity with peaceful coexistence.

There are two aspects to the relationship between the liberal tradition, and the acceptance of diversity and conflict. On one side, the very essence of liberalism is the toleration of diversity

¹⁴⁹ - The challenge facing liberalism is fully discussed in Chapter 1.

¹⁵⁰ - I am using the expression 'hierarchical society' in the same way Rawls does in his Oxford Amnesty Lecture. According to Rawls, hierarchical societies are the alternative to liberal societies, in fact hierarchical societies "are well-ordered and just, often religious in nature and not characterized by the separation of church and state" (Rawls 1993b, p.52).

and the capacity to withstand internal conflict and prosper or improve as a result of it. After all, in the liberal tradition the idea of democracy is defined in terms of pluralism, and you cannot have pluralism without diversity and conflict. To the extent that pluralism is an integral part of the liberal philosophy, diversity and conflict will always be characteristic features of a liberal society.

While diversity and conflict are requisite to liberalism, it is also true that under liberalism the concepts of diversity and conflict acquire specific meanings. Thus we find that diversity must not rule out compatibility, while conflict must be confined to legally recognized and acceptable channels of expression¹⁵¹. Another way of approaching the same conclusion is to observe that in order for a liberal theory to confront the challenge facing liberalism, it must be able to show that diversity and conflict are being tolerated and regulated by a fair system of rules.

In the liberal tradition, theories of the social contract vindicate the attempt to reconcile diversity with peaceful coexistence. In fact the very idea of an agreement is to bring together and reconcile people with different views, that is to say, people whose conceptions of the good are in potential conflict. At the same time, contractualism has also the political function of

¹⁵¹ - These are usually institutional or constitutional channels.

fostering peace and harmony¹⁵², hence while diversity is desired certain types of conflicts must be mitigated.

I think that of all forms of contractualism, the theory of reasonable agreement is best equipped to carry out this double function of seeking the resolution of conflict without crushing the conditions for the development of diversity.

What makes the contractualist theory of reasonable agreement unique respect to other contractualist and liberal theories is the fact that reasonable agreement is imbued in the spirit of scepticism. It is by endorsing the philosophy of scepticism that reasonable agreement is able to respond to the challenge facing liberalism, in fact scepticism encourages agreement between diverse freedoms, while at the same time helping to determine the acceptable limits of diversity and conflict.

I have said before that one of the key notions in the theory of reasonable agreement is the idea of agreement, which implies that a person must be seeking an agreement with others on grounds they could not reject. If by scepticism we understand that frame of mind which challenges us to justify or give reasons for our beliefs, as I have argued in Chapter 1, it follows that endorsing scepticism is a necessary (although not sufficient) condition for

¹⁵² – The most clear exponent of this view is obviously Hobbes, although we find a similar opinion being endorsed by Rawls in the Oxford Amnesty Lecture. According to Rawls, both liberal and hierarchical societies "have a common interest in changing the way in which politics among peoples – war and threats of war – has hitherto been carried out" (Rawls 1993b, p.67), hence Rawls believes that there are the conditions for an agreement between these two types of society.

seeking an agreement on terms that respect the conditions of impartiality. Just as reasonable agreement is about justifying actions to others, scepticism is about justifying one's own beliefs to others.

At the same time, I believe the endorsement of scepticism can help us to establish acceptable limits of diversity and conflict. That is because there is a fundamental difference between the sort of scepticism I am defending (namely, the desire to justify or give reasons for one's beliefs), and the more familiar concept of fallibilism. Basically, the difference is that scepticism applies to our primary or higher conceptions of the good, while fallibilism¹⁵³ applies to our simple conceptions of the good. In other words scepticism deals with the conceptualization of conceptions of the good, hence diversity and conflict between conceptions of the good are tolerated if they are permitted by a system of rules grounded on reasonable agreement, or in other words by the desire to justify or give reasons to others.

So far I have argued that the theory of reasonable agreement, by pitching contractualism at the meta-ethical level, is capable of confronting the challenge facing liberalism. Of course, in order to argue that the contractualist theory of reasonable agreement can be the foundation for a liberal political theory capable of responding to the challenge facing liberalism, it is not sufficient to show

¹⁵³ - I am adopting this term in the way J.S. Mill uses it in On Liberty, namely, fallibilism is the claim that we cannot know everything or alternatively that we cannot be sure that what we believe is the truth.

that at the meta-ethical level reasonable agreement prescribes a different set of moral motivations.

What also needs to be shown is that reasonable agreement can shed some light on some of the key issues at the heart of liberal debates. I have dealt with this question in Part II of the thesis. In these chapters, I have thrust the contractualist theory of reasonable agreement on the mine field of three key concepts in liberal political philosophy: political obligation, social justice and neutrality.

As a theory of political obligation, the idea of reasonable agreement fares better than other approaches, in fact it can justify a state that provides both discretionary goods and presumptive goods, as well as be able to distinguish between a general obligation towards a just state while preserving a prerogative to manifest civil disobedience.

First of all, the idea of agreement, combined with the concept of reasonableness, justifies a duty towards a state that supports the provision of discretionary goods; it is because we are seeking an agreement with others on grounds they could not reject that we are not going to take advantage of our bargaining power, instead we support the provision of discretionary goods. Furthermore, the theory of reasonable agreement respects the right to civil disobedience. This is an important feature of the theory as it is a clear indication of the weight it places on diversity and conflict. After all, civil disobedience is nothing more than a political channel denoting diversity of opinion and reasonable

conflict; if the conflict is not reasonable, we would be dealing with terrorism rather than civil disobedience. Reasonable agreement responds to the challenge facing liberalism by providing a theory of political obligation that recognizes diversity (discretionary goods) and conflict (civil disobedience), hence legitimizing both the state and the boundaries of the political arena.

In terms of social justice, the attempt of reasonable agreement to reconcile the principles of responsibility and compensation is once again a recognition that there are a diversity of people, holding a diversity of legitimate interests and needs, hence only a political theory that aims to reconcile these two principles is attentive to diversity and conflict. The idea of agreement implies that we are not going to support principles of justice simply because they are in our best interest, instead we are seeking an agreement with others. Since those who stand to gain from the principle of responsibility are not those who will gain from the principle of compensation, seeking an agreement necessarily means reconciling these two principles. Furthermore, I believe that the criterion of distribution that reconciles the principles of responsibility and compensation is compatible with Rawls's principles of justice as fairness, in fact the idea of responsibility is reflected in the principle of fair opportunity, while the idea of compensation is reflected in the difference principle.

Finally, contrary to other liberal theories, the concept of neutrality is not refuted by the contractualist theory of

reasonable agreement. Reasonable agreement acknowledges that neutrality and impartiality are closely related. Endorsing neutrality is a recognition that there are a variety of conceptions of the good, hence denying neutrality would be like denying impartiality.

Yet reasonable agreement defends a substantive rather than purely procedural conception of neutrality. In fact neutrality does not apply to every single rules, instead neutrality concerns a system of rules that as a whole is justified by a set of values that no one could reasonably reject. Contractualism in general and reasonable agreement in particular provides the framework in which such values are formulated.

Needless to say, much more work needs to be done before the theory of reasonable agreement can be recognised as providing the moral foundations for the liberal project. My only aim in this thesis was to argue that contractualism is a valid and rewarding approach, which should not be discarded too hastily. Furthermore I have tried to argue that there is more to contractualism than the idea of rational bargaining and mutual advantage, indeed the theory of reasonable agreement is directly antithetical to Gauthier's contractualism.

APPENDIX 'A':

UTILITARIANISM, EGALITARIANISM, AND IMPARTIALITY

Rawls's major critique of utilitarianism, as everyone must know by now, is that utilitarianism does not take seriously the distinction between persons. The force of this argument rests on Rawls's distinction between deontological and teleological theories. According to Rawls, utilitarianism is primarily concerned not with persons, but with their states of affairs, in fact the right act is defined in terms of maximising the good, rather than in terms of equal consideration. Starting from this teleological interpretation of utilitarianism, Rawls concludes that utilitarianism as a moral theory is undesirable because insufficiently egalitarian¹⁵⁴.

In fact Rawls's critique of utilitarianism is not exact. As Kymlicka (1988, 1989a & 1990a) has shown, although Rawls's dissatisfaction with utilitarianism is grounded on the right intuition, he failed to centre the target. Contrary to what Rawls seems to think, utilitarianism does not give priority to the good (maximizing utility) over the right (treating everyone with equal

¹⁵⁴ - See Rawls (1971), p.24.

concern and respect), instead in its more compelling form utilitarianism is a deontological theory, giving priority to the right over the good. As Kymlicka points out, it is the concern with equal consideration that underlies the arguments of Bentham, Sidgwick, and many modern utilitarians like Harsanyi, Griffin, Singer and Hare. According to Kymlicka, utilitarianism has to be examined for what it is, namely, an egalitarian moral theory, working from the same 'egalitarian plateau' as Rawls and Dworkin.

The attractiveness of utilitarianism is that it combines two separate standards: maximising the good and respecting the claim of individuals to equal consideration. Yet, as Kymlicka argues, this is utilitarianism's best asset and worse liability:

Unfortunately, it is incoherent to employ both standards in the same theory. One cannot say that morality is fundamentally about maximizing the good, while also saying that it is fundamentally about respecting the claim of individuals to equal consideration. (Kymlicka 1990a, p.35)

To treat utilitarianism as a purely teleological theory, as Rawls seems to be doing, is to defeat a straw man; the strength of utilitarianism is that its consequentialist character is combined with the core egalitarian intuition of treating people with equal concern and respect. A successful critique of utilitarianism must therefore move away from unfounded teleological charges, and focus on its inadequate conception of equality. In this respect

Kymlicka's views on utilitarianism are an important improvement on Rawls's theory:

while utilitarianism seeks to treat people as equals, it violates many of our intuitions about what it genuinely means to treat people with equal consideration utilitarianism has misinterpreted the ideal of equal consideration for each person's interests, and, as a result, it allows some people to be treated as less than equal, as means to other people's ends. (Kymlicka 1990a, p.36)

The problem with utilitarianism's conception of equality lies in the formula "everyone counting for one, no one for more than one". According to Kymlicka, this is an inadequate account of equality because it assumes that every kind of preference, as long as it yields equal utility, should be given the same weight.

The inadequacy of utilitarian conception of equality originates from its inability to distinguish between different kinds of preferences. External preferences (i.e. prejudice) and selfish preferences (i.e. subsidising expensive tastes on other people's resources) fail to reflect equal consideration, yet these preferences are not ruled out from the utilitarian calculus since doing so would contradict the creed of 'everyone counting for one, no one for more than one'¹⁵⁵. In other words, it is the inability to consider certain preferences as intrinsically unfair that makes

¹⁵⁵ - See Kymlicka (1990a), pp.36-44.

utilitarianism's conception of equality inadequate. An adequate account of equal consideration must distinguish different kinds of preferences, only some of which have legitimate moral weight.

Kymlicka's account of egalitarianism and utilitarianism shows that Rawls's condemnation of utilitarianism is inexact. Considering that Rawls's critique of utilitarianism is the origin of reasonable agreement as a school of political philosophy, it is important to consider the repercussions of Kymlicka's critique of Rawls's views on utilitarianism on the contractualist theory of reasonable agreement. If Rawls's critique of utilitarianism is fallacious, does this undermine the foundations of reasonable agreement? I believe not.

In what follows, I will argue that Rawls's critique of utilitarianism, although inaccurate, played the important role of detaching utilitarianism from the moral perspective of impartiality. In other words, Rawls's critique of utilitarian egalitarianism emancipated the concept of impartiality from the utilitarian clutch. The contractualist theory of reasonable agreement is, first and foremost, a theory of impartiality. Yet before the theory of reasonable agreement could take off, it was necessary to reconceptualize the concept of impartiality in non-utilitarian terms. Rawls's critique of utilitarianism made all this possible.

For the best part of the last twenty years critics of Rawls's theory of justice have been overwhelmingly concerned with his original position, with the unfortunate result that his views on

impartiality have been neglected. It is only recently that this tendency is turning¹⁵⁶. Overlooking Rawls's views on impartiality could have potentially disastrous consequences; it would mean that Rawls's theory of justice amounts to little more than a very sophisticated defence of justice as mutual-advantage.

Of course this is not the case, and Rawls is very explicit about it. At the core of Rawls's theory of justice as fairness is not a desire to come up with a pure procedure grounded on rational-choice theory, but the moral intuition that "moral judgments are, or should be, impartial" (Rawls 1971, p.190). It follows that justice as fairness is not about mutual-advantage, instead it is Rawls's attempt to formalise impartial moral judgments. Similarly Rawls's two principles of justice (the difference principle in particular), are not the result of rational-choice procedure, but the embodiment of a sense of justice grounded on the moral intuition of impartiality¹⁵⁷.

Impartiality is the conceptual key to a correct interpretation of Rawls's theory of justice. Yet, as Rawls was well aware, justice as fairness does not have exclusive rights on impartiality. In 1971 Rawls acknowledged that the utilitarian tradition, especially the classical utilitarian principle of maximizing total utility, also aspires to an impartial moral judgment. Does this mean that justice

¹⁵⁶ – Apart from Scanlon's seminal work in 1982, the works that did most to turn the tide are by Barry (1989b), Beitz (1989) and Nagel (1988).

¹⁵⁷ – For a thorough and compelling interpretation of Rawls's theory of justice along these lines, see Barry (1989b), esp. Parts II and III.

as fairness is a utilitarian doctrine? Obviously not. Rawls explicitly indicates the differences between his theory and utilitarianism, focusing in particular on the different ways in which the concept of impartiality can be interpreted.

In Section 30 of A Theory of Justice¹⁵⁸, Rawls explains the relationship between impartiality, utilitarianism and justice as fairness. He points out that although both utilitarianism and justice as fairness wish to be portrayed as patrons of impartiality, proclaiming their respective principles of justice to be the best interpretation of impartiality, only justice as fairness deserves to be associated with impartiality, since utilitarianism mistakes impartiality with impersonality.

Rawls's aim in Section 30 is to undermine the claim made by utilitarians that the idea of an 'impartial sympathetic spectator' is the correct interpretation of impartiality. According to Rawls this is only one interpretation, and certainly not the correct interpretation of impartiality. The problem with the 'impartial sympathetic spectator' formula is that it relies on the feeling of 'sympathy', and it requires a third party (the 'observer' or 'spectator'). Rawls's idea of impartiality is construed by the veil of ignorance in the original position, whereby sympathy is replaced by mutual-disinterest, and the ideal spectator or observer with the 'litigants' themselves¹⁵⁹.

¹⁵⁸ - 'Classical Utilitarianism, Impartiality and Benevolence'; Rawls (1971), pp.183-192.

¹⁵⁹ - See Rawls (1971), p.190.

In order to understand Rawls's critique of impartiality as sympathetic spectator, it is necessary to recall Rawls's charge that utilitarianism is not sufficiently egalitarian. Earlier I considered some of the problems with Rawls's claim that utilitarianism fails to take seriously the distinction between persons. Notwithstanding these problems, there is an aspect of this critique which must not be neglected - that at the root of Rawls's views is a critique of (classical) utilitarian conception of impartiality; Rawls critique of utilitarian impartiality ought to be read as the core part of his general critique that utilitarianism is insufficiently egalitarian. The central place of impartiality in Rawls's critical analysis of utilitarianism is further evidence of the prominent bearing of impartiality in Rawls's model.

The link between utilitarian conception of impartiality and Rawls's claim that utilitarianism is insufficiently egalitarian is revealed by Rawls with the help of a simple but often over-looked manoeuvre. Rawls blames utilitarianism's insufficient egalitarianism on the logical fallacy of deducing social implications generalizing from single, individual cases. This logical fallacy originates from the endorsement of impartiality as sympathetic spectator. By defining impartiality from the stand-point of a sympathetic observer who responds to the conflicting interests of others as if they were his own, utilitarians mistake impersonality for impartiality, where impersonality implies "the conflation of all desires into one system of desire" (Rawls 1971,

p.188).

According to Rawls, the idea of grounding impartiality on a sympathetic spectator is utilitarianism's major weakness. Rawls's critique of utilitarianism consists in the following line of reasoning: impartiality as sympathetic observer implies impartiality as impersonality, and impartiality as impersonality neglect of separateness of persons, hence the egalitarianism advocated by utilitarianism is inadequate. Having refuted the utilitarian concept of impartiality, Rawls concludes that we ought to abandon utilitarianism and endorse a more egalitarian theory with a different conception of impartiality. Providing a more egalitarianism theory of justice gives justice as fairness the edge over utilitarianism. Justice as fairness is more egalitarian than utilitarianism to the extent that it replaces impartiality as mutual-disinterest for impartiality as sympathetic-spectator.

To the extent that it is also grounded on the concept of impartiality, the theory of reasonable agreement is a direct heir of Rawls's theory of justice. The difference between reasonable agreement and Rawls's theory of justice is that according to the former, impartiality cannot be captured by Rawls's original position, instead impartiality is about the desire to find an agreement on terms that others could not reasonable reject.

BIBLIOGRAPHY

Aarnio, Aulis (1987) The Rational as Reasonable Dordrecht: D. Reidel Publishing Company.

Ackerman, Bruce (1980) Social Justice in the Liberal State New Haven, Conn.: Yale University Press.

Ackerman, Bruce (1983) "What Is Neutral About Neutrality" Ethics Jan. 1983, Vol. 93, No. 2.

Alexander, Larry & Schwarzschild, Maimon (1987), 'Liberalism, Neutrality, and Equality of Welfare vs. Equality of Resources', Philosophy & Public Affairs, Vol. 16, No.1.

Barker, Rodney (1990) Political Legitimacy and the State Oxford: Oxford University Press.

Barry, Brian (1987) "Rationality" in D. Miller (ed.) The Blackwell Encyclopaedia of Political Thought Oxford: Basil Blackwell.

Barry, Brian (1989a) Democracy, Power and Justice: Essays in Political Theory Oxford: Oxford University Press.

Barry, Brian (1989b) A Treatise on Social Justice Vol. I Theories of Justice Berkeley: California University Press.

Barry, Brian (1990a) "The Contractarian Generation" Government and Opposition, Vol. 25 No. 1.

Barry, Brian (1990b) "How Not to Defend Liberal Institutions" British Journal of Political Science Vol. 20, Part 1.

Barry, Brian (1990c) Political Argument: A Reissue with a New Introduction. London: Harvester Wheatsheaf.

Barry, Brian (1991) "Chance, Choice, and Justice" in Liberty and Justice Oxford: Oxford University Press.

Barry, Brian (forthcoming) A Treatise on Social Justice Vol. II The Possibility of Justice Oxford: Oxford University Press.

Beitz, Charles (1979) Political Theory and International Relations Princeton: Princeton University Press.

Beitz, Charles (1989) Political Equality Princeton: Princeton University Press.

Bellamy, Richard (1989) "Defining Liberalism: Neutralist, Ethical or Political?" in R. Bellamy (ed.) Liberalism and Recent Legal and Social Philosophy Franz Steiner Verlag Weisbaden GMBH, Stuttgart.

Bellamy, Richard (1994) "The Anti-Poll Tax Non-Payment Campaign and Liberal Concepts of Political Obligation" Government and Opposition, 1994, Vol. 29, No. 1.

Berlin, Isaiah (1969) "John Stuart Mill and the Ends of Life", in Four Essays on Liberty, Oxford: Oxford University Press.

Bobbio, Norberto (1987) The Future of Democracy Cambridge: Polity Press.

Botwinick, Aryeh (1990) Skepticism and Political Participation Philadelphia: Temple University Press.

Campbell, Tom (1988) Justice Basingstoke: Macmillan.

Caney, Simon (1991), "Consequentialist Defences of Liberal Neutrality", The Philosophical Quarterly, Vol. 41, No.165.

Cohen, G.A. (1992) "Incentives, Inequality, and Community" in The Tanner Lectures on Human Values Vol.13, Salt Lake City: University of Utah Press.

Cohen-Almagor, Raphael (1993) "Harm-Principle, Offence Principle, and the Skokie Affair" Political Studies Vol.41, No.3.

Copi, Irving H. (1986 7th ed.) Introduction to Logic, New York: Macmillan.

Daniels, Norman (1975) "Introduction" in N.Daniels (ed.) Reading Rawls Oxford: Blackwell.

Davidson, D. (1963) "Actions, Reasons, and Causes" Journal of Philosophy, Vol.60, No.23.

Davidson, D. (1985) "Replies to Essays I-IX", in B.Vermazen & M.B.Hintikka (eds.) Essays on Davidson. Actions and Events Oxford: Clarendon Press.

Del Vecchio, Giorgio (1952) Justice: An Historical and Philosophical Essay Edinburgh: Edinburgh University Press.

De Marneffe, Peter (1990) "Liberalism, Liberty, and Neutrality", Philosophy & Public Affairs Vol.19, No.3.

Dworkin, Ronald (1978) "Liberalism" in S.Hampshire (ed.) Public and Private Morality Cambridge: Cambridge University Press.

Dworkin, Ronald (1981a) "Equality of What? Part I: Equality of Welfare" Philosophy & Public Affairs Vol. 10, No.3.

Dworkin, Ronald (1981b) "Equality of What? Part II: Equality of Resources" Philosophy & Public Affairs Vol.10, No.4.

Dworkin, Ronald (1983a) "What Liberalism Isn't", The New York Review of Books, Jan.20, 1983.

Dworkin, Ronald (1983b) "Why Liberals Should Care About Equality", New York Review of Books, Feb.3, 1983.

Dworkin, Ronald (1985) "Liberalism" in A Matter of Principle Oxford: Clarendon Press.

Dworkin, Ronald (1990) "Foundations of Liberal Equality" The Tanner Lectures on Human Value Vol.XI, University of Utah Press: Salt Lake City.

Eatwell, Roger (1993) "Ideologies: Approaches and Trends" in R.Eatwell & A.Wright (eds.) Contemporary Political Ideologies London: Pinter Publishers.

Elster, Jon (1985) "The Nature and Scope of Rational-Choice Explanation", in E.LePore and B.P.McLaughlin (ed.) Actions and Events Oxford: Basil Blackwell.

Elster, Jon (1989) Solomonic Judgments Cambridge: Cambridge University Press.

Finnis, John (1980) Natural Law and Natural Rights Oxford: Oxford University Press.

Fishkin, James (1984) Beyond Subjective Morality New Haven, Conn.: Yale University Press.

Fishkin, James (1992) The Dialogue of Justice: Toward a Self-Reflective Society New Haven: Yale University Press.

Frankfurt, Harry (1971) "Freedom of the Will and the Concept of a Person" Journal of Philosophy Vol.68.

Franklin, James (1991) "Healthy Scepticism" Philosophy July 1991, Vol.66, No.257.

Freeman, Samuel (1990) "Reason and Agreement in Social Contract Views" Philosophy and Public Affairs, Spring 1990, Vol.19, No.2.

Freeman, Samuel (1991) "Contractualism, Moral Motivation, and Practical Reason" The Journal of Philosophy Vol.LXXXVIII, No.6.

Gardbaum, Stephen A. (1992) "Law, Politics, and the Claims of Community" Michigan Law Review Vol.90.

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

Gibbard, Allan (1990) Wise Choices, Apt Feelings: a Theory of Normative Judgment Oxford: Clarendon Press.

Gibbard, Allan (1991) "Constructing Justice" Philosophy & Public Affairs Vol.20, No.3.

Gibson, John (1990) The Politics and Economics of the Poll Tax: Mrs. Thatcher's Downfall Warley: Emas.

Gilbert, Margaret (1993) "Agreement, Coercion, and Obligation" Ethics Vol.103, No.4.

Goodin, Robert E. (1993) "Equal Rationality and Initial Endowments" in D.Gauthier & R.Sugden (eds.) Rationality, Justice and the Social Contract Hempstead: Harvester Wheatsheaf.

Goodin, Robert & Reeve, Andrew (1989) "Liberalism and Neutrality" in R.E.Goodin & A.Reeve (eds.) Liberal Neutrality London: Routledge.

Goodwin, Barbara (1992) Using Political Ideas, third edition, Chichester: John Wiley & Sons.

Gray, John (1989) Liberalisms: Essays in Political Philosophy London: Routledge.

Grayling, A.C. (1985) The Refutation of Scepticism London: Duckworth.

Grayling, A.C. (1991) "Testing Notions of Knowledge" TLS September 6 1991.

Green, Leslie (1990) The Authority of the State Oxford: Clarendon.

Hamlin, Alan. & Pettit, Philip (eds.) (1989) The Good Polity Oxford: Basil Blackwell.

Hamlin, Alan (1989) "Liberty, Contract and the State" in A.Hamlin & P.Pettit (eds.) The Good Polity Oxford: Basil Blackwell.

Hampton, J. (1980) "Contracts and Choices: Does Rawls Have a Social Contract Theory?" Journal of Philosophy Vol. 77.

Hampton, J. (1991) "Two Faces of Contractarian Thought" in P.Vallentyne (ed.) Contractarianism and Rational Choice New York: Cambridge University Press.

Hampshire, Stuart (1983) Morality and Conflict Oxford: Basil Blackwell.

Hampsher-Monk, Iain (1992) A History of Modern Political Thought Oxford: Basil Blackwell.

Hardin, Russell (1991) book review of Beitz's Political Equality, in Political Theory Vol.19, No.4.

Hare, Richard M. (1981) Moral Thinking Oxford: Clarendon Press.

Hart, H.L.A. (1955) "Are There ANY Natural Rights?" The

Philosophical Review, Vol.LXIV, No.2.

Hobbes, Thomas (1651) Leviathan Edited by C.B.Macpherson (1968), Harmondsworth: Penguin.

Hookway, Christopher (1990) Scepticism London: Routledge.

Jones, Peter (1989a) 'Liberalism, Belief and Doubt' in R. Bellamy (ed.) Liberalism and Recent Legal and Social Philosophy

Jones, Peter (1989b) 'The Ideal of the Neutral State', in R.E.Goodin & A.Reeve (eds.) Liberal Neutrality London: Routledge.

Kant, Immanuel (1991, 2nd ed.) "On the Common Saying: 'This May Be True in Theory, But It Does Not Apply in Practice'" in Kant: Political Writings, Hans Reiss (ed.) Cambridge: Cambridge University Press.

Klosko, George (1987) "The Principle of Fairness and Political Obligation" Ethics, Vol.97, No.2.

Klosko, George (1992) The Principle of Fairness and Political Obligation, Lanham, MD: Rowman & Littlefield.

Kukathas, Chandran & Pettit, Philip (1990) Rawls: A Theory of Justice and its Critics Cambridge: Polity 1990.

Kymlicka, Will (1988) "Rawls on Teleology and Deontology" Philosophy and Public Affairs Vol. 17, No.3.

Kymlicka, Will (1989a) Liberalism, Community and Culture Oxford: Oxford University Press.

Kymlicka, Will (1989b) "Liberal Individualism and Liberal Neutrality" Ethics Vol.99, No.4.

Kymlicka, Will (1990a) Contemporary Political Philosophy, Oxford: Oxford University Press.

Kymlicka, Will (1990b) "Two Theories of Justice" Inquiry, Vol.33, No.1.

Kymlicka, Will (1991) "The Social Contract Tradition" in P.Singer (ed.) A Companion to Ethics Oxford: Basil Blackwell.

Kymlicka, Will (1993) "Community" in R.E.Goodin & P.Pettit (eds.) A Companion to Contemporary Political Philosophy Oxford: Basil Blackwell.

Larmore, Charles E. (1987) Patterns of Moral Complexity Cambridge: Cambridge University Press.

Lehning, Percy (1991) "Liberalism and Capabilities: Theories of Justice and the Neutral State" Social Justice Research.

Lehning, Percy (1993) "Right Constraints?" in D.Gauthier & R.Sugden (eds.) Rationality, Justice and the Social Contract Hempstead: Harvester Wheatsheaf.

Lessnoff, Michael (1986) Social Contract Basingstoke: Macmillan.

Lessnoff, Michael (ed.) (1990) "Introduction: Social Contract" in Social Contract Theory Oxford: Basil Blackwell 1990.

Letwin, William (ed.) (1983) Against Equality London: Macmillan.

MacKinnon, Kenneth A (1990) "The 'Reasonable Man' as an Impartial Spectator" in T.Campbell (ed.) Law and Enlightenment in Britain, Aberdeen: Aberdeen University Press.

Mill, John Stuart (1972) Utilitarianism, On Liberty and Considerations on Representative Government, H.B. Acton (ed.), London: Dent.

Minogue, Kenneth (1989) The Egalitarian Conceit: True and False

Equalities London: Centre for Policy Studies.

Montefiore, Alan (1975) Neutrality and Impartiality Cambridge: Cambridge University Press.

Moya, Carlos J. (1990) The Philosophy of Action: an Introduction Cambridge: Polity Press.

Murphy, Jeffrie G. (1988) "Afterword: Constitutionalism, Moral Skepticism, and Religious Belief", in Alan S. Rosenbaum (ed.) Constitutionalism: the Philosophical Dimension, Westport, Conn.: Greenwood Press.

Nagel, Thomas (1979) Mortal Questions Cambridge: Cambridge University Press.

Nagel, Thomas (1987) "Moral Conflict and Political Legitimacy" Philosophy and Public Affairs Summer 1987, Vol.16, No.3

Nagel, Thomas (1988) 'The Foundations of Impartiality' in Seanor & Foton (eds.) Hare and Critics Oxford: Clarendon Press.

Nagel, Thomas (1991) Equality and Partiality Oxford: Oxford University Press.

Nielsen, Kai (1992) "Conceptions of Justice", in M. Hawkesworth & M. Kogan (eds.) Encyclopedia of Government and Politics, Vol. I London & New York: Routledge.

Nozick, Robert (1984) Anarchy, State, and Utopia Oxford: Basil Blackwell.

Pateman, Carole (1985) The Problem of Political Obligation London: Polity Press.

Parekh, Bhikhu (1993) "A Misconceived Discourse on Political Obligation" Political Studies Vol.41, No.2.

Perelman, Chaim (1963) The Idea of Justice and the Problem of Argument London: Routledge & Kegan Paul.

Perelman, Chaim (1979) The New Rhetoric and the Humanities Dordrecht: D. Reidel Publishing Company

Perry, Thomas D. (1976) Moral Reasoning and Truth Oxford: Oxford University Press.

Plant, Raymond (1991) Modern Political Thought Oxford: Basil Blackwell.

Putnam, Hilary (1981) Reason, Truth, and History Cambridge: Cambridge University Press.

Rawls, John (1971) A Theory of Justice Oxford: Oxford University Press.

Rawls, John (1980) "Kantian Constructivism in Moral Theory" The Journal of Philosophy Vol. 77, No.9.

Rawls, John (1982) "Social Unity and Primary Goods" in A. Sen and B. Williams (eds.) Utilitarianism and Beyond Cambridge: Cambridge University Press.

Rawls, John (1985) "Justice as Fairness: Political not Metaphysical" Philosophy & Public Affairs Vol.14, No.3

Rawls, John (1987), "The Idea of an Overlapping Consensus", Oxford Journal of Legal Studies, Spring 1987, Vol.7, No.1.

Rawls, John (1988), "The Priority of Right and Ideas of the Good", Philosophy & Public Affairs, Fall 1988, Vol.17, No.4.

Rawls, John (1989) "The Domain of the Political and Overlapping Consensus", New York University Law Review, May 1989, Vol.64, No.2.

Rawls, John (1993a) Political Liberalism New York: Columbia University Press.

Rawls, John (1993b) "The Law of Peoples" in S.L. Hurley and S. Shute (eds.) On Human Rights, New York: Basic Books.

Rawls, John (MS) Justice as Fairness: A Brief Restatement Harvard University 1989.

Raz, Joseph (1986) The Morality of Freedom, Oxford: Clarendon Press.

Raz, Joseph (1989) "Liberalism, Skepticism, and Democracy", Iowa Law Review, May 1989, Vol. 74, No.4.

Richardson, H. (1990) "The Problem of Liberalism and the Good" in B.Douglass, G.Mara, H.Richardson (eds.) Liberalism and the Good, London & New York: Routledge.

Ripstein, Arthur (1994) "Equality, Luck, and Responsibility" Philosophy & Public Affairs Vol.23, No.1.

Sandel, Michael (1982) Liberalism and the Limits of Justice Cambridge: Cambridge University Press.

Scanlon, Thomas (1975) "Preference and Urgency" The Journal of Philosophy Vol.82, No.19.

Scanlon, Thomas (1977) "Liberty, Contract, and Contribution" in G.Dworkin, G.Bermant, & P.Brown (eds.) Markets and Morals New York & London: John Wiley & Sons.

Scanlon, Thomas (1982) "Contractualism and Utilitarianism" in A. Sen & B. Williams (eds) Utilitarianism and Beyond, Cambridge: Cambridge University Press.

Scanlon, Thomas (1986) "Equality of Resources and Equality of Welfare: a Forces Marriage?" Ethics, Vol. 97, No.1.

Scanlon, Thomas (1988a) "Levels of Moral Thinking", in Seanor & Foton (eds.) Hare and Critics Oxford: Clarendon Press.

Scanlon, Thomas (1988b) "The Significance of Choice" The Tanner Lectures on Human Values Vol.8, Salt lake City: University of Utah Press.

Scanlon, Thomas (1992) "The Aims and Authority of Moral Theory" Oxford Journal of Legal Studies Vol. 12, No. 1.

Scheffler, Samuel (1992) "Responsibility, Reactive Attitudes, and Liberalism in Philosophy and Politics" Philosophy & Public Affairs Vol.21, No.4.

Scruton, Roger (1984 2nd ed.) The Meaning of Conservatism Basingstoke: Macmillan.

Sher, George (1987) Desert Princeton, NJ: Princeton University Press.

Sibley, W.M. (1953) "The Rational Versus the Reasonable" The Philosophical Review, Vol.LXII, No.4.

Simmons, Alan J. (1979) Moral Principles and Political Obligations Princeton: Princeton University Press.

Tocqueville, Alexis (1840) Democracy in America Vol.II, New York: Vintage Books Edition, 1990.

Van Parijs, Philippe (1991) "Why Surfers Should be Fed: The Liberal Case for an Unconditional Basic Income" Philosophy & Public Affairs, Vol.20, No.2.

Waldron, Jeremy (1987) "Theoretical Foundations of Liberalism" The Philosophical Quarterly, Vol.37, No.147.

Waldron, Jeremy (1989) "Rights in Conflict" Ethics April 1989, Vol.99.

Waldron, Jeremy (1993) "Special Ties and Natural Duties" Philosophy & Public Affairs, Vol.22, No.1.

Waldron, Jeremy (1994) "The Edges of Life" book review of R.Dworkin's Life's Dominion, in London Review of Books 12 May 1994, Vol.16, No.9.

Walzer, Michael (1983) Spheres of Justice Oxford: Basil Blackwell.

Weale, Albert (1993) "Justice, Social Union and the Separateness of Persons" in D.Gauthier & R.Sugden (eds.) Rationality, Justice and the Social Contract Hempstead: Harvester Wheatsheaf.

Williams, Bernard (1985) Ethics and the Limits of Philosophy London: Fontana.

Wolf, Naomi (1990) The Beauty Myth London: Chatto & Windus.