

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

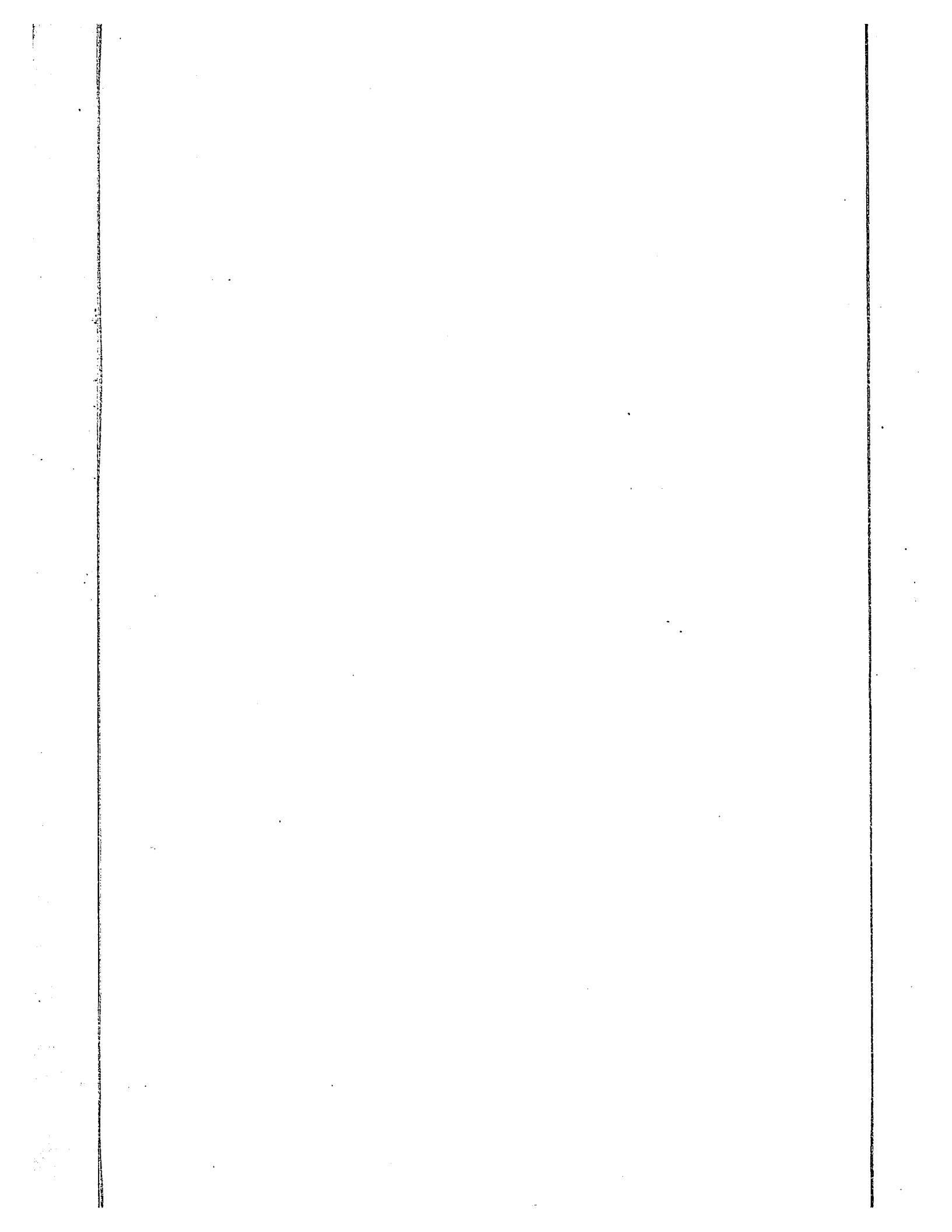
The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]



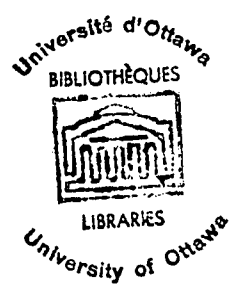
Csp-

"THE 'GOOD' IN JOHN RAWLS' CONTRACTARIAN THEORY OF JUSTICE"

By

Joseph Tunde Erumevba

Thesis submitted to the School of
Graduate Studies of the University
of Ottawa in partial fulfillment
of the requirements for the Ph. D.
degree in Philosophy



OTTAWA, CANADA, SEPTEMBER 1977.

UMI Number: DC52345

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

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

UMI[®]

UMI Microform DC52345
Copyright 2007 by ProQuest LLC
All rights reserved. This microform edition 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

DEDICATION

TO THE ETERNAL MEMORY OF MY MOTHER VERONICA WHO DEPARTED
FROM THIS WORLD JUST AS THIS STUDY WAS BEING CONCLUDED.

ACKNOWLEDGMENTS

In a sense, no student can rightly claim to be the sole author of a thesis. He owes so many debts to those from whom he has learned (both formally and informally) and from whom he has received inspiration and encouragement. The author of this thesis is no exception. Special thanks are due to my supervisor, Professor Kai Nielsen of the University of Calgary, Calgary, Alberta. Professor Nielsen gave me unlimited access not only to his many published writings on Rawls, but to an extensive collection of works in contemporary ethical theory. Also, special thanks go to the staff of the faculty of Philosophy of the University of Ottawa. Each time I was in difficulty (this was not infrequent.) I knew there was always someone at the faculty with a useful idea. I am particularly grateful to my good friend Hyacinth Joy Allen who typed all of my first draft. Finally, gracious and special thanks to my friend Yvrose Vilfort. Her words of encouragement provided me with great upliftment when the going was rough.

TABLE OF CONTENTS

	<u>Page</u>
Acknowledgments	
<u>Chapter:</u>	
INTRODUCTION	1
OVERVIEW	1
JUSTICE AS FAIRNESS	13
THE CHALLENGE OF MORAL SCEPTICISM	29
PLAN & METHOD OF APPROACH	47
I. THE ORIGINAL POSITION	54
PREVIEW: RAWLS' IDEA OF THE SOCIAL CONTRACT	54
A THEORY OF ANALYSIS AND CRITICISM	59
ISOLATING THE MOST REASONABLE INTERPRETATION OF THE INITIAL SITUATION	63
A THEORY OF MORAL JUSTIFICATION	76
THE DERIVATION OF THE TWO PRINCIPLES OF JUSTICE AS FAIRNESS	82
THE CHOICE OF MORAL PRINCIPLES FROM BEHIND A VEIL OF IGNORANCE	109
SUMMARY AND CONCLUSION	120
II. RAWLS' MORAL EPISTEMOLOGY	126
MORAL PHILOSOPHY: AN ATTEMPT TO EXPLAIN OUR MORAL CAPACITY	130
OUR CONSIDERED JUDGMENTS IN REFLECTIVE EQUILIBRIUM	137
III. THE PRINCIPLES OF JUSTICE AS FAIRNESS	159

THE PRINCIPLES OF GREATEST EQUAL LIBERTY	160
RESTRICTION OF LIBERTY FOR THE SAKE OF LIBERTY	160
THE MAXIMIN CRITERION	177
CONCLUSION	202
IV. THE PRIMARY GOOD	205
PREVIEW: THE ROLE OF THE GOOD IN JUSTICE AS FAIRNESS	205
THE INDEX OF PRIMARY GOODS	215
V. RATIONALITY IN THE THIN THEORY OF THE GOOD	237
THE PRINCIPLES OF RATIONAL CHOICE	253
DELIBERATIVE RATIONALITY	255
THE ARISTOTELIAN PRINCIPLE	268
VI. THE FULL THEORY OF THE GOOD	287
A GOOD PERSON OR A PERSON OF MORAL WORTH	288
THE DESIRE TO ACT JUSTLY	302
THE MORALITY OF AUTHORITY	305
THE MORALITY OF PRINCIPLES	317
VII. THE GOOD OF THE SENSE OF JUSTICE	345
THE QUESTION OF CONGRUENCE	345
VIII. SOME PROPOSALS TOWARDS A RECONSTRUCTION OF RAWLS' ACCOUNT.....	400
BEYOND THE RATIONALITY OF THE PRUDENT MAXIMIZER	407
TOWARD AN ENRICHED CONCEPTION OF RATIONALITY	427

THE POSSIBILITY OF A CONTRACT	441
CONCLUSION: A LESSON FROM THE PLIGHT OF RAWLS' THEORY	452
SELECTED BIBLIOGRAPHY	460

INTRODUCTION

OVERVIEW

(1)
John Rawls' A Theory of Justice has been widely extolled as an epic work in moral and political philosophy in several scholarly circles. Brian Barry refers to it "as a work of great significance for moral and political philosophy ... a work that anyone in the future who proposes to deal with any of the topics it touches must first come to terms with if he expects the scholarly community to take him seriously"⁽²⁾.

Barry's observation alludes to a very important consideration vis-a-vis any appraisal of the worth of A Theory of Justice. The vast amount of critical reviews which has appeared ever since its many years of gestation, leaves one wondering at first, as to what is so important about this work. One becomes even more puzzled when one notes the fact that the general thrust of these reviews is that Rawls' theory does not work and that his arguments are unsound. Granted that this is indeed the case, why, one might ask, does a defective theory attract so much attention?

1 John Rawls, A Theory of Justice, Harvard University Press, 1972, pp.xv - 608.

2 Brian Barry, The Liberal Theory of Justice, Clarendon Press, 1973, p.ix.

Barry's observation might be quite helpful in dispelling this perplexity. The significance of Rawls' work may lie not so much in the correctness of his substantive principles of justice as it does in his revival of the architectonic of moral theory.⁽³⁾ Even if this approach turns out to be unsuccessful in the end, it is nonetheless impressive. Kai Nielsen outlines some of its more obvious features as follows:

1. In a post-Wittgensteinian and post-Austinian intellectual environment, Rawls, without compromising the standards of conceptual sophistication and analytical rigour that we have come to appreciate, has shown us that, and how we can do systematic moral and political philosophy in a way which is clearly neither normatively neutral nor moralistic. He has shown us that good moral and political philosophy need not be piecemeal analysis and he has shown us, in Hampshire's words, that moral philosophy need not be an "intellectually barren subject in which no major discoveries were to be expected and no new lines of inquiry could be opened".
2. In focussing, as he does, on principles of justice and a contractarian account, Rawls examines closely the place of reason in morality and articulates clearly principles of rational choice. Whatever may be the defects of his account of rationality, his situating it as he does and his stress on the need to clearly articulate and defend principles of rational choice, is a major advance in moral philosophy.

3 Kant represents the idea of architectonic as "the act of constructing systems". He argues that since "systematic unity is what first raises ordinary knowledge to the rank of science, that is, makes a system out of mere aggregate of knowledge, architectonic is the doctrine of the scientific in our knowledge". (Kant, Critique of Pure Reason, trans. Norman Kemp Smith, London: Macmillan W., 1961, p.653.

3. Rawls, not by frontal attacks on moral scepticism and not by attempts to defuse subjectivism, but in the very example of his inquiry, has effectively challenged (or so it would seem) what Hampshire has called the "sceptical orthodoxy" which has developed from an ever present, but typically minority position, to at least a dominant tradition in twentieth century moral philosophy. Rawls has not attempted to show that moral scepticism or subjectivism is in some way incoherent; rather, he has attempted to show that it is unnecessary. That is to say, he has shown, or at least attempted to show, how moral beliefs have a rational foundation, and how questions of substance can be asked and answered within moral philosophy. He attempts to show us the kind of principles we must be reasoning in accordance with when a moral stand is a reasonable one. And, by their very articulation, he has shown us that they are available to us.
4. Again, in a way that invites comparison with Sidgwick, Rawls deeply and carefully focusses on the fundamental question of method concerning: (a) how we can reasonably test substantive moral and political claims, (b) how we can assess rival moral theories and (c) how we can reasonably use a moral theory in assessing particular moral claims. (4)

Thus, A Theory of Justice can be seen to draw an instructive comparison with The Critique of Pure Reason. Rawls himself says, "I must disclaim originality for the views I put forward. The leading ones are classical and well-known. My intention has been to organize them into a general framework by using certain simplifying devices so that their full force can be appreciated. My ambitions for the book will be completely realized if it enables one to see more clearly the chief structural features of the alternative conception of justice that is implicit in the contract tradition and points the way to its further elaboration". (Rawls, op. cit., p.viii). It remains to be seen in the course of this study whether Rawls can claim success even in such a modest objective.

4. Kai Nielsen, "A Critical Notice of John Rawls' A Theory of Justice in Second Order: An African Journal of Philosophy, Vol. II, No.2, July 1973, pp.101-121.

It is important to note that the very framework offered by Rawls for the consideration of the problems of morality turns out to be an effective basis for criticizing his account of justice.

If indeed Rawls is able to deliver on these claims, he would be advancing the course of moral philosophy in a very significant way. He would, as it were, be granting the entire discipline a new lease of life. It is useful in a preliminary way to expound further on the issue of moral scepticism in general and, in particular, an element of it, the challenge of amoralism, in view of the general thrust of this study.⁽⁵⁾ To provide an appropriate background for this, we will first examine some general considerations that led to the emergence of A Theory of Justice.

WHY A NEW THEORY OF JUSTICE

The very title of Rawls' book stirs within the reader a certain degree of curiosity as to the need for a new theory of justice. Is there any justification for a new theory of justice? This question becomes very pertinent if one takes into account the vast amount of literature that has been

5 Since the dawn of the modern era, especially since the publication of G.E. Moore's Principia, (1903), one of the central issues that has persistently beset the minds of moral philosophers is the question of providing adequate justification for moral first principles: a justification that would constrain all reflective persons, including the moral sceptic to acknowledge the intrinsic desirability of acting from the moral point of view. Rawls takes considerable interest in this question. At least the approach he follows in his theory of justice takes it into account. Hence one of the main interests of this study is to determine what contribution he has made toward the resolution of this question, with particular reference to the challenge of amoralism.

published on the notion of justice dating back to the time of Plato. If after all this time moral philosophy has not been able to evolve principles for the realization of the good life or just society, why should this area of study continue to be treated with respect? Besides, if Rawls intends his account as a fulfillment of the philosophers' hope for an adequate conceptualization of a just society, why has he not chosen a stronger title such as "The Theory of Justice" instead of "A Theory of Justice"?

These questions might seem trivial at first glance. But, in the light of the persistent attack on the validity and relevance of moral philosophy since the dawn of the modern era, it has become necessary for the philosopher who wants to speculate on the moral phenomenon to justify himself almost before he gets a hearing. Such a justification becomes all the more necessary especially if he attempts to go beyond the analysis of moral concepts, and tries to deal with substantive issues in morality.

Perhaps the desire to attain an Archimedian point for assessing the whole of morality, which Rawls proudly announces at the beginning of his book, does suggest an air of novelty in his theory. It is doubtful, all the same, whether Rawls actually sees A Theory of Justice as attempting something radically new in ethical theory. His remark concerning his architectonic plan does not suggest this.

Along with his modern predecessors, notably Kant, Rawls does not conceive the task of the moral philosopher as that of "inventing" or "discovering" new ethical principles. Rather, Rawls' aim, as he states in the preface of his book, is to present "an alternative systematic account of justice that is superior to the dominant utilitarianism of the tradition", based on the point of view of social contract.⁽⁶⁾ He observes that utilitarianism in one form or another has been "the predominant systematic theory" in modern philosophy.⁽⁷⁾ One reason for this, he contends, is that the doctrine "has been espoused by a long line of brilliant writers who have built up a body of thought truly impressive in its scope and refinement".⁽⁸⁾ However, we are wont to forget, Rawls claims, "that the great utilitarians, Hume and Adam Smith, Bentham and Mill were social theorists and economists of the first rank, and the moral doctrine they worked out was framed to meet the needs of their wider interests and to fit a comprehensive scheme".⁽⁹⁾ One of the most distinguishing traits about these philosophers is their basic approach to moral philosophy. In contrast to the trend which still prevails among

6 Rawls, op. cit., p.viii.

7 Ibid.

8 Ibid.

9 Ibid.

many philosophers of the Anglo-Saxon tradition, these philosophers do not narrow the task of moral philosophy to a descriptive analysis of the logical behaviour of moral terms or sentences (meta-ethics). For them, moral philosophy is an integral part of critical and systematic social theory. Hence as moral philosophers, they had no inhibition in making contingent assumptions, and when it suits them they also appealed to general facts.

Rawls is in agreement with the classical utilitarians on the role of moral philosophy to the extent that moral philosophy is broadly conceived by the latter as embracing social theory.⁽¹⁰⁾ Nevertheless, he rejects the substantive doctrine of utilitarianism in all its various forms. This rejection need not be seen as a surprise. The doctrine of utilitarianism has been criticized at every stage of its development all through the history of Western thought. But as Rawls points out, most of these criticisms have often been made "on a much narrower front. They pointed out the obscurities

10 Peter Singer challenges this claim. He argues in his paper "Sidgwick and Reflective Equilibrium" that Rawls' methodological procedure in appealing to the notion of reflective equilibrium cannot be associated with Sidgwick. Sidgwick, he claims, does not use this method and the method itself must be wrong. There is a fundamental difference, no doubt, between Sidgwick and Rawls, especially in what they take to be the basic principles of morality. But this fact by itself does not refute the claim that both of them see the nature of ethics, or its place in social theory, in an identical way. The disagreement between them is on the principles or reasons explaining the "facts" of morality and not on what

of the principles of utility and noted the apparent incongruities between many of its implications and our moral sentiments⁽¹¹⁾. The upshot of such criticisms is that they do not leave us "a workable and systematic moral conception to oppose it"⁽¹²⁾. Hence, we are left in a position where we must necessarily choose between utilitarianism and intuitionism. And, as Rawls observes, in most cases we settle for "a variant of the utility principle circumscribed and restricted in certain ad hoc ways by intuitionistic constraints"⁽¹³⁾.

Thus the purpose of Rawls' project is to correct this situation. In A Theory of Justice, Rawls wants to develop a systematic alternative to utilitarianism: he seeks to construct a critical social theory that rests on a non-utilitarian foundation. He undertakes this task by generalizing and carrying "to a higher order of abstraction the traditional theory of the social contract as represented by Locke, Rousseau, and Kant"⁽¹⁴⁾.

should constitute moral "facts". This point will become evident when we come to examine Rawls' moral methodology in Chapter Two. (Peter Singer, "Sidgwick and Reflective Equilibrium", *The Monist*, 58, (July, 1974), pp.490-517.

11 Rawls, op. cit., p.viii.

12 Ibid.

13 Ibid.

It should be noted that Rawls uses the term "intuitionism" here in a more restricted sense. He uses it to refer to the doctrine "that there is an irreducible family of first principles which have to be weighed against one another by asking ourselves which balance in our considered judgments, is the most just.

One might readily agree with Rawls on the need for a systematic alternative theory to utilitarianism. Apart from the familiar criticism that the utility principle does not offer an adequate theory of moral justification, there is also the question about the adequacy of the concept of right it offers. Following the utility principle, rightness consists in whatever produces the greatest good for the greatest number. It would appear at first that in this formulation, the right and the good are actually defined simultaneously. On close inspection however, the right (the greatest) emerges merely as adjectival to the good and dependent on it. But the concept of greatest (or the idea of maximization) is obviously not a necessary feature of the concept of right, which suggests that the right can in fact be independent of the good. But this does not seem to be the case in the context of the utility principle. Within the utilitarian tradition the right is linked with the good, and the good is treated as prior to the right. Hence the determination of what is right or obligatory would then depend upon a prior determination of the good.

Once we attain a certain level of generality, the intuitionist maintains that there exists no higher-order constructive criteria for determining the proper emphasis of the competing principles of justice". Rawls, Ibid., p.34.

To be sure, the utilitarian's way of linking the right and the good may seem natural enough if ethics were indeed to begin with the good and then seeks a definition of the right. Given such a starting point, the basis for thinking that this association is natural lies simply in the fact that we then tend to think of the good as something that should be maximized, more of a good being better than less. But the direct consequence of such a move is that one is both deprived of an independent concept of right as well as the freedom to implement a conception of good, if such a conception is incompatible with the dominant conception of good in a particular society. A majoritarian democracy would be a useful example here. In such a system, the popular conception of the good could easily override that of the minority even if the minority conception is more enlightened. Consider the thorny issue of the abolition of slavery in the Congressional debates on the draft constitution of the American independence. The abolition clause had to be deleted from the draft constitution before it was passed. Thus by allowing the good to set the stage for the definition of the right, it seems possible that some putative wrongs can be defined as right, if the majority considers such wrongs to be good in a particular society. Perhaps there are some important ways of meeting these problems. The framework of this study does not involve examining the complexities of utilitarianism. Rather

the intent of these remarks is simply to show some of the general considerations which led to Rawls' A Theory of Justice.

Now, assuming that Rawls does indeed have a prima facie case for rejecting utilitarianism, and for subsequently wanting to develop an alternative theory, the question yet remains: why does he think that it is possible to return to a tradition which has been so thoroughly discredited in the eyes of many contemporary philosophers in order to remedy at least the putative inadequacies of the utilitarian moral point of view? Tradition as used here refers specifically to the notion of an original social contract. This notion has been vigorously criticized as far back as the time of Hume. As Joe H. Hicks points out, the contract theory has been persistently repudiated as "a fiction of Enlightenment political philosophy: empirically unsupported, based on a model of detached rationality, blind to the forces of psychology, sociology, and history, and finally lacking teeth morally and politically"⁽¹⁵⁾. In that case, is Rawls not undercutting his own grounds in selecting the contractarian approach as a way of expounding his conception of justice? It is of course the privilege of any theorist to adopt whatever method or approach he might consider suitable in the development of his insights. However,

15 Joe H. Hicks "Philosophers' Contracts and the Law", Ethics, Vol. 85, No. 1, Oct., 1974.

in making a positive return to a tradition whose reasoning has been widely rejected as both delusive and futile, Rawls must do at least one of two things. Apart from defending the substantive claims of his conception of justice, he must be able to show that the charges against the contract tradition are either ill-founded or that his own version of the social contract can effectively by-pass them. Rawls tries to do the latter. While he evokes the notion of an original social contract, especially in its Kantian form, he attempts as it were to give it new flesh. As Hicks further points out, Rawls' version of the social contract introduces "methodological clarity, plausible empirical assumptions, an elaborate socio-economic model of rationality, historical relevance, and moral determinateness"⁽¹⁶⁾. Thus, it would appear that Rawls has actually attempted a substantial revision of the traditional doctrine of the social contract. And if this is the case then his claim that he has merely generalized and carried "to a higher order of abstraction the traditional theory of the social contract as represented by Locke, Rousseau and Kant"⁽¹⁷⁾, would seem to be very much a question of intellectual

16 Ibid.

17 Rawls, op. cit., p.viii.

modesty. However, it remains to be seen in the course of this study whether his version of the social contract can in actual fact undermine the difficulties traditionally associated with the social contract doctrine. Meanwhile, we will examine some of the general ideas in Rawls' theory.

JUSTICE AS FAIRNESS

Rawls' attempt to define an alternative conception to the utilitarian conception of justice is the result of a theory of distributive justice which he labels as "Justice as Fairness". Using the notion of fairness both as the fundamental concept of justice and as an independent basis for the principles of right prior to the consideration of any "major" conception of good, Rawls seeks to evolve a formula for regulating the distribution of all benefits and burdens, rights and duties in a society. His central assumption, as Stuart Hampshire points out, is that "an injustice is an arbitrary inequality in the distribution of good things and primarily in the

18 The term "major" is quite significant here. Even though Rawls maintains the priority of the right over the good, he still uses a minimal conception of good that does actually precede and influence in some sense his theory of justice. He has to rely upon what he calls "primary goods" and a "thin theory of the good" in order to be able to formulate any principles of justice at all. Thus, in a sense, Rawls might yet concede that right is impossibly abstract without a prior knowledge of good. But such knowledge need not compromise an independent conception of right if a proper procedure is followed. Rawls shows how this could be done by assuming a thin theory of the good in his original position.

distribution of liberty. Inequality of rights and liberty, and inequalities in the distribution of good things generally, are arbitrary from a moral point of view if they cannot be shown to produce a benefit in which all the least fortunate (19) have a preponderant share". In order that any distribution might be acceptable from the moral point of view, it must not only be just, it must be fair as well. (20) To understand the radical nature of Rawls' conception of fairness, and to be able to distinguish it from the concept of justice, it is necessary to refer to a thought-experiment which he calls the "original position".

The original position is Rawls' own interpretation of the hypothetical initial situation that must be on hand in order to derive the principles of justice. It corresponds roughly to the idea of "the state of nature" in traditional contract theories. Rawls wants to use it to represent a moral standpoint wherein individuals are assumed to be behind a "veil of ignorance" that shields certain important personal information from them. They do not know who they are, their place in society, and the nature of their individual conceptions

19 Stuart Hampshire, "What is The Just Society?", The New York Review of Books, xviii (February 24, 1972) p.34.

20 Ordinarily one might think that the concept of justice is virtually synonymous with fairness. This is not the case as far as Rawls is concerned. While he sees fairness as being fundamental to the concept of justice, the notion of fairness is yet distinct from that of justice.

of the good. From such a purely hypothetical situation, individuals are to choose the principles of justice under identical constraints, the most important of which is the lack of self-knowledge, including their position in society and their particular conception of the good. However, Rawls allows the contractors certain general information in order to ensure the emergence of a workable conception of justice. Hence, they are presumed to know the general facts about human society. They understand political affairs and the principles of economics; and, they know the basis of social organization and the laws of human psychology. (21)

Also, the parties in the original position are presumed to be rational to the extent that they are capable of taking effective means towards their desired ends (whatever the latter might be). They also know that, whatever rational plan of life each of them might subsequently want to pursue, there are certain "rock bottom" goods (primary goods) which

21 This assumption has evoked some criticism in several reviews on Rawls' theory. For example, Robert Paul Wolff maintains that there are powerful reasons for at least doubting the cognitive possibility of such a general knowledge. (See Robert Paul Wolff, Understanding Rawls, Princeton Univ. Press, 1976, pp.119-132.) It appears that several of Rawls' conclusions are indeed based on questionable sociological and psychological assumptions. For instance, the assertion that the parties in the original position know the "laws of human psychology" seem to derive from an easy assumption that knowledge of human personality, like knowledge of bodies in motion, is atemporal and impersonal. While followers of Piaget and the like might share such an assumption, it is definitely not acceptable to a Freudian. From a Freudian standpoint, we cannot possibly understand human

must be on hand in order to realize their plans of life. Hence, they also know that having more of these primary goods is more desirable than having less of them. In regard to the question of interpersonal attitudes, the parties are to be seen as mutually self-interested. They do not take interest in the affairs of one another, though they are neither egoists nor are they envious of one another.

Perhaps the most crucial constraint in the original position is the elimination of the parties' individual conceptions of good. Rawls introduces this constraint in view of the fact that among any group of individuals, there are always several conceptions of good. This fact, as Rawls sees it, can only promote disagreement over the principles of justice. The original position is essentially a simplifying device that is required to define fair principles of justice on the prior basis of a fair set of conditions. Put in very simple terms, the idea is that a fair result can only

psychology except in relation to some particular experience. Thus the parties in the original position could not really understand anything about human beings since they are deprived of all personal information. Perhaps Rawls might respond to this by saying that the framework of the original position is satisfactory, and we can simply plug in alternative social and psychological theories and see what comes out. Hence, the question for him is not whether people like the persons in the original position could really exist but what we here and now think about things if we abstract from certain features of our own situation. Such a response does seem adequate. However, it remains to be seen whether such a weaker rendering of the original position can effectively fulfill the role Rawls designs for it.

be the outcome of a choice procedure that is fairly designed. By eliminating the parties' individual conceptions of the good, Rawls wants to show that our nature as moral persons does not depend entirely on our aims or our conceptions of the good. Rather, in keeping with the doctrine of the priority of right, our nature as moral persons is revealed by "the principles that we would acknowledge to govern the back-ground conditions under which these aims are to be pursued"⁽²²⁾. And, as it emerges from A Theory of Justice, such principles must include the priority of equal liberty for all persons under ideal conditions, and in less than ideal situations in which inequalities are inevitable, then such inequalities must primarily favour the worst-off members of the society.

The general idea behind justice as fairness does seem to make a strong initial appeal. To begin with, Rawls' attempt to treat the concept of right as prior to the concept of good has the prospect of yielding an independent criterion for the concept of right. Rawls' appeal here consists primarily in the principle that the good, whatever it is, must be susceptible to criteria other than actual human desires even if such desires happen to be those of the majority. It might so happen in particular instances that the desires of the majority do in fact correspond to the dominant conception of good.

22 Rawls, op. cit., p.560.

Nevertheless, it need not be right to allow such desires to prevail unless they also coincide with the morally good. For a thing to be morally good, it must, according to Rawls, be consistent with the principles of right.

Also, the elimination of certain particular information in the original position does seem to evoke some intuitive support. For instance, if it is indeed the case that the factual aspects of the personal lives of the contractors would inhibit their capacity to respond impartially, then it is only reasonable that such information be suppressed. Should Rawls admit the facts characterizing one's personal life as a basis for developing the principles of justice; then he might as well end up with what in Husserlian language might be termed "a prejudice of the world". But his very notion of fairness requires a correction of the moral arbitrariness of the world: arbitrariness which is often manifested in the manner in which society and natural endowment have favoured certain individuals with more advantageous starting points. It is Rawls' firm belief that to take the natural endowment of nature and society as a point of departure for adjudicating issues of justice -- issues of what is due to persons in regard to the common good -- is to make a virtue out of arbitrariness. Unless good reason can be given to show that a person actually merits or deserves such assets or defects, then any use of them is arbitrary or unfair and hence unjust.

The original contractors who desire and see themselves as equals would not wish to see some of them benefiting from the advantages that customarily go with native or social endowments. The results that flow from using them were determined on an unfair basis. (23)

Thus in a move that draws comparison with Husserlian reduction the original position tries to move the deliberators away from natural and social contingencies; contingencies which are morally irrelevant. They cannot arrive at a conception of pure morality or justice if their considerations begin with the natural and social worlds. (24)

By suppressing the individual conceptions of the good in the original position, Rawls assumes that rational and self-interested persons would know that they ought not jeopardize their eventual disposition or hypothetical freedoms by evoking a particular conception of good in the original position. If they are rational, and if they do not as yet know what their actual conceptions of good will turn out to be when the veil of ignorance is lifted, then they should be

23 Perhaps Rawls does not actually intend to say that inequalities are intrinsically unjust, but that their influence should be of less importance in considering the question of justice. However, he does seem to be over-stating the case. While the outcomes of the natural and social lotteries of life, to use his terms, are indeed unmerited it is not so evident that those so endowed are necessarily acting unjustly by benefiting from them. As David Gauthier rightly asks, why is it unjust for someone to benefit from his natural talents? There are several prominent people in our society who owe their prominence to their natural talents. Perhaps Rawls' own chair of

reluctant in acceding to the risk of accepting any conception of good as a premise in considering what principles of justice to adopt. In view of the constraint of strict compliance, they know that once they have made their choice of principles then they must stick by them.

Now despite this initial appeal, Rawls' justice as fairness still raises some difficulties which are not so obviously resolvable. Perhaps the most important difficulty is the precise status or relation of the notion of fairness to the original position. While Rawls is inclined to view a theory of justice as part of the theory of rational choice, such that justice will be decided by what it would be rational to choose under the conditions of the original position, it is not so evident whether these conditions can actually yield the notion of fairness. (Recall the distinction that Rawls maintains between justice and fairness. Justice is not fairness. Both are logically distinct even though the latter is fundamental to the concept of justice.) From Rawls' characterization

philosophy at Harvard is largely due to his natural talents as a philosophical genius. Now why is it unjust for him or anyone else to benefit from these talents? Also, it is not evident why the deliberators of the original position would not agree that someone should not benefit from his talents. It is true that they see themselves as equals. But we must not forget that they are also self-interested persons who are neither envious of one another nor even take interest in the affairs of others. (David Gauthier, "Justice and Natural Endowment: Toward a Critique of Rawls' Ideological Framework", Social Theory and Practice, 3, (1974), pp.3-26.

of fairness, as well as the conditions of the original position, neither of these can be entirely or even primarily a result of rationality (at least not a value-free conception of rationality). One does not arrive at these conditions through a discursive reasoning. In fact by assuming these conditions, what Rawls seems to want to show above anything else is that the requirements of fairness demand that we should blot out or strip ourselves of the contingent donations of nature and nurture and then confront ourselves as rational moral beings who are free from individual and social prejudice, and who are capable of appreciating the intrinsic worth of morality or justice (fairness). But such a demand does not appeal strictly to rationality, at least not rationality in a morally neutral sense. It appeals to certain values outside of it. Now how does one then justify the demand of fairness?

It is necessary to recognize at this point two separate types of arguments at the original position, one argument concerned with justice as such, and the other with fairness. The argument for justice consists essentially in determining what it is strictly rational to choose given the

24 What Rawls is attempting to do here also draws instructive comparison with Kant's project of isolating "pure moral philosophy" from practical anthropology in the Foundations of the Metaphysics of morals. See G.J. Warnock's "KANT AND ANTHROPOLOGY" in Nature & conduct, ed. R.S. Peters, pp. 36-46. Robert Paul Wolff, op. cit., pp. 112-116.

circumstances and constraints of the original position. The argument for fairness, on the other hand, does not so depend on the nature of rationality. Rather, to establish the notion of fairness, Rawls relies on our moral sentiments which he standardly refers to as moral "intuitions". These are the moral convictions which are generally fixed and firm for all or nearly all men. They are, as it were, humanity's firm moral convictions about fairness. Thus one fundamental trait of justice as fairness is Rawls' attempt to delineate fairness by appealing to what he believes are humanity's firm moral convictions about fairness, and then proceeding to derive the principles of justice from the conditions that define fairness as exemplified in the original position. (25) By virtue of fairness an individual is expected to divest himself of the natural and social contingencies of his own life, contingencies which ordinarily dictate his plan of life and his conception of the

25 There are two key questions that should be kept in mind here. First, in view of the fact of such an obvious diversity in people's beliefs and values, how can one maintain in a nontrivial way that there are such things as "humanity's firm moral convictions"? Several of Rawls' critics, notably Peter Singer, Steven Lukes, Bernard Williams, Kai Nielsen et al have challenged this claim vigorously. They argue that the moral convictions that Rawls appeals to are culturally determined. They are the moral beliefs typically shared by Western liberal democrats. This question will be examined carefully in Chapter Two. Secondly, assuming that one can indeed specify some moral beliefs that are commonly held by all of humanity, and, that an appeal to these beliefs is necessary in order to arrive at the notion of fairness, why does Rawls still need a social contract or the device of the original position? How necessary is the contract argument within the framework of

good. Given these conditions of fairness (which also are the conditions of the original position), the individual, Rawls believes, if he is rational, will demand that other deliberators on justice, together with himself should refrain from appealing to any distinctive conception of the good in trying to choose the principles of justice.

At this point some of the fundamental puzzles about Rawls' theory begin to surface. Let us suppose that a person, given the circumstances of the original position, will find it rational to choose the principles of justice as fairness. What is the moral significance of such a prudential choice? Why are the principles chosen under these conditions necessarily of fairness? Perhaps one should take into account here the two types of arguments mentioned in the preceding paragraph (the argument of justice and the argument of fairness). But even then, what grounds do we have to believe that the conditions of the original position adequately represent those of fairness? Besides, does the very appeal to moral intuitions not expose justice as fairness to a sort of moral relativism, be it a relativism of a rather subtle type? Of course Rawls would have his reader believe that this appeal is quite

Rawls' project of trying to show the notion of fairness as being fundamental to the concept of justice? If the requirements of fairness are not the same as those of ordinary rational choice, what is the relevance of the device of the original position whose main strategy is designed to show the conditions of rational choice? Such an issue will continue to emerge in one form or another all through this study. It is doubtful whether there is a basis for a satisfactory solution to it in Rawls' account.

consistent with and indeed comparable to the standard procedure in linguistics. The linguistic theorist, he says, relies on the intuitions of a native speaker about what is an acceptable sentence of his language. In like manner, we should check the outcomes of particular moral and social theories against our moral intuitions, and seek to attain a state of "reflective equilibrium".

This analogy has been sharply criticized by several of Rawls' critics. It is said not to work mainly because the linguistic intuitions of native speakers are decisive as regards grammar in a way in which moral intuitions are not, for moral principles or theories. For example, Thomas Nagel notes that whatever the ordinary native speakers of English would agree on as English can be accepted without question as English. But whatever ordinary men agree in condemning need not be necessarily wrong. Thus, whereas the intrinsic plausibility of an ethical theory can impel a change in our moral convictions, it is the linguist who must change his theory to conform to the intuitions of the native speakers should a change be
(26)
desirable at all.

26 Thomas Nagel, "Rawls on Justice", in Reading Rawls, ed., Norman Daniels, New York, 1974, p.2.

While this remark is obviously correct, Rawls' analogy is still defensible if one takes into account what kind of theory, he thinks a theory of justice is supposed to be. As mentioned earlier, the task of the moral philosopher, as conceived by Rawls, is to describe and to explain our moral sentiments (sense of justice). To accomplish this task the moral philosopher formulates a set of principles which, when conjoined to people's beliefs and knowledge of the circumstances, would lead them to make the judgments with their supporting reasons which they do actually make. From this standpoint, a moral theory does seem analogous to a theory of linguistics in describing and explaining the sense of grammaticalness for sentences of one's native language. Now, while one might be able to sustain Rawls' analogy by consideration such as this, there are still two important questions that must be answered. (1) If the two principles of justice as fairness are only a descriptive-explanatory hypothesis, as his conception of the nature of moral theory seems to imply, what is the relevance of the intricate procedure of their contractarian derivation from the original position? The point at issue here is not the analytic role of the original position, but its justificatory role. A contractarian strategy may still be useful as a way of discovering what moral judgments people do in fact make in the context of a descriptive-explanatory theory. But this is not

just what Rawls is after. He wants the choice of the principles in the original position to be a justification of these principles. He does not want to take our actual considered judgments about justice for granted. Rather, the principles of justice as fairness are to prescribe what our opinion and judgments about particular instances ought to be. But then how can justice as fairness be a descriptive-explanatory hypothesis as well as a prescriptive-justificatory hypothesis all at the same time? It appears that Rawls has not been able to recognize the distinction between these two types of theories. His attempt to conflate both of them into one has seriously diminished the effectiveness of his moral methodology. (2) Let us suppose that a theory of justice is indeed a descriptive-explanatory hypothesis, as Rawls sees it. In which case, it takes its "facts", namely, people's moral intuitions (their common-sense convictions and judgments of what is just and unjust) as it finds them. Now is there sufficient agreement among the considered moral judgments of different individuals for them to be derivable from the same set of principles? Are there such fixed intuitions about justice and morality that all men share in common regardless of such contingencies as race, colour and creed? As mentioned earlier, Rawls' procedure in appealing to people's actual judgments in particular cases need not commit him to a subjectivist viewpoint, if one takes the

descriptive-explanatory character of his theory into account. Hence, he might still be able to turn out an objective theory of justice even when he does rely ultimately on people's actual judgments in particular cases, as the truth condition for the validity of his principles. However, it is not Rawls' methodological procedure itself that is in question at this point. The issue is the possibility of attaining what he refers to as the "Archimedian point" through his methodological procedure. Rawls wants to be able to subsume the majority, if not all, of the considered moral judgments that people do make under the principles of justice as fairness. If the principles of justice as fairness are well-founded, then we should have on hand, a set of principles with which we could regulate and determine the justness of basic social institutions and all moralities. Now if it is the case, as it appears to be, that there is a diversity in people's considered moral judgments, do we not then need various sets of principles or theories of justice to account for all the disagreements in people's considered moral judgments about what counts as just? And, also, would the validity of justice as fairness not then be relative only to the moral convictions of a particular group of people (the Western liberal democrats)? It does not appear that Rawls has been able to provide any convincing argument why this question should not be answered in the affirmative. The issue will be discussed

in some detail in Chapter Two.

THE PROBLEM TO BE INVESTIGATED

The questions raised in the preceding paragraphs indicate the general dimension of this study. The central focus of the inquiry is the justificatory force of Rawls' theory of justice. As remarked earlier, while the very emergence of A Theory of Justice may be said to have granted moral philosophy a new lease of life, it remains to be seen whether the extended arguments for the insights contained in this work can withstand critical probing without collapsing into various errors or trivialities. And, if it is indeed the case, as suggested by several critics, that the theory is plagued by serious difficulties, is there a way of dissolving some of these without any essential modification of either its principles or its framework? This is the problem that the thesis undertakes to study.

The study will be done with particular reference to Rawls' theory of "Goodness as Rationality". By examining this theory in the context of his overall account of justice, we will try to determine to what extent Rawls has been able to overcome some of the central worries in modern ethical theory, especially the challenge of moral scepticism. As Nielsen rightly points out, even though Rawls does not engage in frontal attacks on moral scepticism or any of the varieties

of subjectivism in ethics, the very example of his inquiry is a challenge to such ways of viewing morality. By constructing and defending a descriptive ethical theory which makes truth claims, and which is based on what he claims are objective principles, Rawls implicitly attempts to undermine the thesis of moral scepticism and subjectivism by doing, or at least attempting to do what such accounts claim cannot be done.⁽²⁷⁾ Thus in this study we will try to determine whether Rawls' account of the good is a coherent one, squaring adequately with his overall account of morality and whether it, and his theory generally, is immune to the challenge of moral scepticism. A brief remark on some aspects of this challenge may help articulate more fully the general approach of this study.

THE CHALLENGE OF MORAL SCEPTICISM

The challenge of radical scepticism remains one of the central problems of philosophy. In the preface to the Critique of Pure Reason, Kant tells us that "it still remains a scandal to philosophy and human reason in general, that the existence of things outside us -- must be accepted merely on faith, and that if anyone thinks it good to doubt their existence, we are unable to counter his doubts by any satisfactory proof".⁽²⁸⁾ The

27 Kai Nielsen, op. cit., pp.102-3.

28 Immanuel Kant, Critique of Pure Reason, trans. N.K. Smith. p.34.

essence of the sceptical challenge is the demand for a complete justification of the system of our beliefs. Although Kant's comment refers specifically to our common sense belief in the existence of an external world, the target of the sceptical attack is much wider than this. The sceptic calls into question the entire system of our rational beliefs. While he acknowledges the fact that some of these beliefs may be partially justifiable insofar as they are grounded on other beliefs, he maintains that ultimately all our beliefs rest on a non-rational commitment. They lack a rationally supportable foundation. Hence, they are in the final analysis arbitrary.

In the context of morality, the challenge of scepticism takes several different forms. It includes various forms of non-cognitivism, ethical subjectivism and relativism. However, the remarks in this section will focus on two forms of this challenge which seem to have direct relation to certain features in Rawls' theory. These are the claims of amoralism and ethical subjectivism. First we examine the amoralist claim.

The challenge of amoralism focuses mainly on the problem of adducing adequate justification for the moral point of view. Here the sceptic poses -- though this is not the only question he poses -- the unusual question "Why Be Moral?". The point of his challenge might properly be viewed in the context of the old moral dilemma between the concerns of the self (rational self-interest) and the concerns for the interests,

needs and welfare of others. Why must a person promote the interests of others when it is not in his self-interest to do so?

Although there are fundamental differences among the various traditions in ethics on such questions as ethical first principles and the method of arriving at them, most writers in moral philosophy would still agree that for any principles to count as a moral principle or an ethical standard, it must at the very least accommodate the interests of others. Ethical egoism is usually considered inadequate either as a normative or meta-ethical theory. Kai Nielsen argues this point rather convincingly in his refutation of ethical egoism. He says :

Universal categorical ethical egoism will not do as a meta-ethical theory purporting to analyze what is meant when people say something is morally good or obligatory; and if it is offered as a radical normative ethic it likewise gets into intolerable paradoxes. In order to remain intelligible, egoism must be put forth as individual and not as a universal egoism. But, as such, it also fails to meet the minimum conditions necessary for something to count as 'morality', as that word is ordinarily and intelligibly used -- To count as an ethical doctrine, then an ethical egoism of ends must be understood as claiming that everyone ought to seek his own good as an end and consider the good of others only when this would in his judgment further his own good -- such a standard could not be moral for we have moral standards to impartially adjudicate the conflicting interests of individuals and groups; but if each individual's own rational self-interest is taken as the standard, in reality we have no standard by which to adjudicate these conflicting interests. The very raison d'etre of morality has been frustrated. Thus self-interest, no matter how enlightened; cannot be our standard of moral appraisal. (29)

Nielsen's position is fair enough and the sceptic going the route of the amoralist might even agree with him without relinquishing his challenge. If indeed an agent has already decided to act morally, then his principle must at least accommodate the welfare of others. This is one of the minimal conditions on which the entire fabric of morality rests. However, while accepting this fact and understanding about morality, the sceptic may yet come back with the question: why should an individual commit himself to the moral point of view at all? It is important to note here that the sceptic or the amoralist is not really challenging the practical usefulness of social morality. He might even be the first to agree that our society, or indeed any society, can hardly survive without some collective morality. But his attack is not so much against the rationale of a "collective morality", as against "individual morality". Why should an individual be moral at all times even when he could get away undetected and the society not suffering any real harm if he decides not to act morally? Why should anyone at all voluntarily accept the constraints of morality when it is not in his personal (or class) interest to do so?

29 Kai Nielsen, "Why Should I Be Moral?" Readings in Ethical Theory, ed. Wilfrid Sellars & John Hospers, New York, 1970, pp.748-9.

Some philosophers have attempted to defuse this aspect of moral scepticism by trying to show that its demand is illegitimate. The question "why be moral?", they argue, is both unwarranted and unjustified. Hare's hypothetical case of the Nazi Jew is a typical example in this line of attack. Should a Nazi-minded Jew declare that "all Jews ought to be exterminated", this proposition, according to Hare, is not only universalizable, it is by its very form actually universalized. And if it is indeed a universal proposition, then the proponent must acknowledge its applicability to himself even if it turns out that he is of semitic origin himself. But the latter condition, Hare argues, only goes to show that the proposition is both unwarranted and unjustified in the first instance, because no one except a fanatic would voluntarily accept its consequences. Of course one might ask: why not be a fanatic? To the latter question one could only reply that the term "fanatic" as it functions in ordinary parlance is used to describe those whose behaviour is less than rational. Of course we need to bear in mind here that the term "fanatic" is not purely descriptive or value-free.

This reductio type argument seems, at first sight, quite effective against the amoralist. If he is going to wave his axe indiscriminately against the whole of morality, he might end up cutting his own throat. Nonetheless, it yet seems that Hare's hypothetical case begs a fundamental issue.

The crucial point of the sceptical challenge, as indicated earlier, is the very legitimacy of the moral point of view. And this is precisely the question that Hare's argument seems to beg. Surely if the Nazi Jew has already committed himself to the moral point of view and he is willing to play the moral language-game, he cannot will, without contradicting himself, that the consequences of his judgment be not applicable to himself. He would be contradicting himself, if he asserts that all Jews "ought" (ought in a moral sense) to be executed, while at the same time asserting that he himself, a Jew, should not (should in a moral sense) be executed. But why must he adopt the moral point of view? Why must he play the moral language-game, thereby giving his "ought" the force of a moral ought?

It would seem that the anti-semite Jew would have successfully countered Hare's argument by simply dissociating himself from the moral point of view and insisting that his "ought" is a non-moral ought. For the intelligibility of his demand it is sufficient that he harbours a genuine hatred for all Jews the world over, and be personally determined to ensure that they are murdered if he were to have his way. Of course, Hare might still wonder whether this man would yet press his demand after he has come to discover that he is of semitic origin himself. Even with such a discovery the Nazi Jew need not abandon his position. All he needs to do is to

impose a constraint on his proposition such as "all Jews except his family, friends and himself ought to be exterminated". The mere fact that he is a Jew himself does not necessarily abate his hatred for Jews. He hates all Jews, except himself, and perhaps his associates as well, and he thinks they ought (non-moral ought) to be exterminated.

Perhaps one might be able to impugne the sceptical attack on other grounds but, as far as the intelligibility of the position itself is concerned, it does not seem to be so apparently irrational, at least not if one takes the term "rational" in a morally neutral sense. One important factor that must be noted here is the crucial distinction between a moral "ought" and a non-moral "ought". These are two different uses. As Nielsen rightly observes, it makes sense to ask "why should people be moral?" and "why should I be moral?", provided the "should" in the two questions is not construed as a moral "should". The questions become trivial and absurd, such as "why are all round things circular?", only if the "should" is construed in a moral sense. There are several uses of "should" and "ought" which are not moral uses. (30)

Thus, though the questions "why should people be moral?" and "why should I be moral?" may indeed sound unusual, they are both intelligible and legitimate. They portray at least a

30 Ibid. p.748.

putatively legitimate demand by the sceptic for a satisfactory justification of morality. If it is indeed the case that an outstanding characteristic of morality requires substantial personal sacrifices on the part of the moral agent, then it seems quite legitimate to inquire whether such sacrifices or the entire moral point of view is really necessary and justifiable. As Kurt Baier notes in the preface to his book The Moral Point of View, one of the four logical features of moral judgments is that "there must be good reasons why any and every moral agent should do the morally right thing rather than the opposite"⁽³¹⁾. It is the possibility of such good reasons that the moral sceptic, going the route of amoralism, calls into question.

Now assuming that the demand of the amoralist is indeed coherent and legitimate, what implication does it have for Rawls? In what way precisely does the challenge of amoralism pose a threat to justice as fairness?

As noted earlier, the very nature of Rawls' theory is a counter attack on moral scepticism. By viewing a theory of justice as part of a theory of rational choice whose principles have been clearly articulated, Rawls is (at least) attempting to show that moral beliefs have a rational foundation and that questions of substance can both be asked and answered within

31 Kurt Baier, The Moral Point of View, London, 1974, p.v.

moral philosophy. The principles emerging from Rawls' theory of justice as fairness and his principles of rational choice together constitute the cluster of principles that must be on hand for a person's moral stand to be considered a reasonable one. Hence, contrary to the claim of the sceptic, Rawls maintains that there is a link between morality and rationality.

Rawls' account of Goodness as Rationality is particularly significant vis-à-vis the sceptical challenge. Among other things, the theory purports to offer a guarantee of "some reasonably secure sense of what things rational individuals -- whatever else they want -- will desire as prerequisites for carrying out their plans of life"⁽³²⁾. Also in what might be construed as an important answer to the question "Why Be Moral?", Rawls' theory of the good offers to show that there are at least certain conditions in which "being a good person is a good thing for that person"⁽³³⁾. The members of a well-ordered society, if they are rational, will affirm their sense of justice and demand that others do likewise. Now, if Rawls can indeed deliver on these claims, he would have defused an important aspect of the sceptical challenge and thus overcome one of the most resilient difficulties in moral philosophy.

32 Rawls, op. cit., p.397.

33 Ibid.

The other form of moral scepticism which appears to raise some challenge for Rawls' theory is ethical subjectivism. As indicated earlier, some critics of Rawls argue rather forcefully that his methodological procedure, in particular his appeal to the "considered moral judgments" of individuals as a way of testing the validity of his moral principles, commits his theory to radical subjectivism. To be able to determine whether or not Rawls is in fact vulnerable to this charge, it is necessary to ascertain just what is at issue. What precisely is involved in the claim of ethical subjectivism and how coherent is this claim?

To answer this question, one needs to examine the following three statements, each of which, in its different way, expresses a view that moral convictions or judgments are "merely subjective".

1. Moral judgments are subjective because they merely describe or express the attitudes of the utterer.
2. Moral judgments are subjective because, unlike scientific statements, they cannot be made or justified independently of the attitudes of some particular human being or groups of human beings, they are only matters of individual opinion.
3. Moral judgments cannot be objective because there are no moral facts; there are only the sorts of facts that science or common observation can discover, and the values that men place on those facts. (34)

34 Bernard Williams, Morality: An Introduction to Ethics, New York, 1972, p.13; also see Kai Nielsen, "Does ETHICAL SUBJECTIVISM HAVE A COHERENT FORM?", Philosophy and Phenomenological Research, Vol., 35(1), 1974, p.93.

These statements appear very close to one another. In fact the proponents of ethical subjectivism tend to use them rather interchangeably in defending their viewpoint. They all articulate the familiar, or at least putative, distinction between fact and value. Moral judgments, they claim, are not of the realm of fact; they belong to the realm of value. Since values are in themselves subjective, moral judgments must also be subjective. However, the statements are significantly different in their emphasis. The first statement emphasizes the logical behavior of moral judgments (what moral judgments do). The second statement is primarily concerned with issues of epistemology: moral judgments are not the sort of judgments that can be justified, they are unlike scientific statements. The third statement appears quite capable of collapsing into the second. Yet it does project something distinctive. It expresses a metaphysical viewpoint about the world. It denies the existence of a moral order "out there". What exists in the world are only the sorts of things or "facts" that science can investigate.

The issue involved here has been of great concern to modern moral philosophy. If there is indeed a distinction between fact and value what implication does it have for moral remarks? One important way in which moral philosophers have tried to respond to this question is to acknowledge the distinction as being essential to the very nature of morality,

while maintaining at the same time that it does not deter the objective validity of moral statements. ⁽³⁵⁾ How adequate is this response? This will become evident as we examine each of the statements cited above.

The first statement is either false or trivially harmless. It is false if the subjectivist claim is to the effect that moral judgments only state the attitudes held by their utterers. For this claim to be true, it must be the case that moral statements are no more than mere autobiographical claims. But this is not the case. The statement "X is morally good" or "X is the right thing to do" is not just replaceable without the loss of anything, by the statement "I approve of X" or "some individuals or groups approve of X". Besides, if moral judgments are reducible to mere autobiographical claims, there would be no point in people disagreeing over moral claims. But people do in fact disagree over moral claims. Someone might say that "X is the right (morally good) thing to do", while another person claims the contrary. Nor is it a conceptual anomaly for one to claim that "I like doing X, I approve of X, but X is not the morally right thing to do", or that "People approve of X but they should not". But how can these statements be meaningful if moral claims only state the attitudes of their utterers?

35 Bernard Williams, op. cit., pp.14-15. Also see Hare's Language of Morals.

It would appear, therefore, that since moral claims are capable of generating conflict, they cannot just be interpreted as autobiographical claims. They do not merely describe the attitudes of their utterers.

But this is not to say that moral statements do not have the function of expressing the attitude of the speaker. To be sure, a moral judgment does indicate what the attitude of its utterer is. But this claim in itself, as Bernard Williams rightly points out, is incontestable and harmless. In itself, it comes to no more than saying that a man who makes a moral judgment can (if that judgment is sincere) be said to be expressing his attitude to a certain moral issue, and this does not support any distinctively subjectivist view of moral judgments. A man who sincerely makes a factual judgment may be said to be expressing his belief about a factual issue, but that does not support a subjectivist view of factual judgments. (36) Perhaps the subjectivist might insist that a moral judgment can be considered objective only if it refers to a reality which does not depend for its existence on any human attitudes, feelings or intentions. But even then his claim would be vacuous since it is utterly impossible to conceive of such a reality.

36 Ibid., P. 16

To be anything more than a triviality, the subjectivist claim must be interpreted as saying that moral judgments can only express the attitudes of the utterers and nothing more. Such attitudes, unlike those a person expresses in a factual judgment, do not raise the question of right or wrong. But this claim is plainly false. As mentioned earlier, it is a generally recognized feature about morality, that moral judgments, unlike utterances of taste and the like can be "right" or "wrong". When a person makes a moral claim, he is normally expected to advance good reasons for its justification. Now does the idea or possibility of adducing good reasons not indicate that it is assumed that there exists a background of some commonly accepted moral attitudes? For individuals to have a meaningful argument or disagreement about moral claims, they must at least have a common background of moral agreement: they must share some moral attitudes in the background which are not in dispute. ⁽³⁷⁾ And if this is the case, the claim that a moral utterance merely describes the attitude of the speaker must be false.

The subjectivist might still be able to respond to this argument. He might concede the fact that the attitude a person expresses in a moral judgment is indeed not the kind of attitude he would have towards some brand of ice-cream or

37 Ibid, p.18.

automobile. The subjectivist might also acknowledge the fact that when moral claims are made, there is a general demand for an exchange of good reasons. But he could still go on to argue that these facts only go to show that the morality of a person or society is to some extent general and systematic and that general attitudes can, to a certain extent, be applied to particular and less general cases. In which case individuals might be able to argue within the framework of some commonly accepted moral attitudes or principles. But outside such a framework there is no way of showing any particular moral view to be right or wrong. In other words, morality is ultimately relative. Even though moral agents by and large believe that they make objective moral claims (claims that require exchange of good reasons), they cannot provide conclusive justification for such claims especially if they involve questions of ethical first principles. There is no objective method for resolving disagreements over ultimate moral principles.

The subjectivist argument has apparently taken a new dimension. Its focus has now shifted from the logical or linguistic status of moral utterances (what moral utterances are or do). What the subjectivist is expressing here is an epistemological view about moral utterances. And this leads naturally to the consideration of the second statement of the subjectivist thesis.

The second statement of the subjectivist thesis is a far more serious threat to the assumed objectivity of moral judgments. It would appear that if the latter statement is true, the possibility of a rational normative ethics will be very much in doubt. The kind of objectivity that the subjectivist is demanding here is one in virtue of which the truth or falsity of a moral statement does not depend on the peculiarities of the person who makes the statement or the culture to which he belongs. Rather the truth or falsity of a moral claim should be ascertainable by any rational person who is
(38)
apprised of the facts.

To be sure one might be able to raise difficulties about this demand. For instance one could ask whether this kind of objectivity is even a legitimate demand in the social or behavioral sciences. Apart from the physical sciences, can any statement (or principle) be objectively true or false in a sense that does not depend on the peculiarities of the person who makes the statement or the culture to which he belongs? Now if, as it seems, the social and behavioral sciences do not require such an 'absolute' objectivity to construct valid and

38 It is instructive to think of Westermarck's subjectivist view of ethics here. In his Ethical Relativity, Westermarck argues that, if moral judgments state objective truths, there must be considerations by which their truth can be established. But he says that typical ethical theories, including hedonism, utilitarianism, evolutionary ethics, rationalism, and the various accounts of a special "moral faculty", are quite unable to defend their basic principles. (J.L.Mackie, "Westermarck", The Encyclopedia of Phil., Vol.8, ed., Paul Edwards, p.285).

objective theories, why is it so crucial for ethical theory? Perhaps some consideration along this line may tend to mitigate the force of the subjectivist demand. However there is still some genuine puzzle about this demand. The subjectivist can effectively counter the suggestion that the social and behavioral sciences do not require the type of objectivity that he demands of ethics. He might say that the social and behavioral sciences, unlike ethical theory, do not prescribe what "ought" to be. Their activity is descriptive and not normative. Moreover, whatever conclusions are reached in these sciences there are concrete observable data with which one could check for truth or validity. Besides, the subjectivist might admit that our ordinary moral judgments do make a claim to objectivity. But his point is that such a claim by itself does not show that any moral judgments have objective validity. Moral judgments, for what they are, only result from the "objectivizing" of moral emotions, "this being just one example of 'a very general tendency to assign objectivity to our subjective experience'⁽³⁹⁾. If this point is valid, it would certainly undermine one of the arguments advanced earlier against the first statement of ethical subjectivism (naive subjectivism). It would undermine the attempt to support ethical objectivism by appealing to the meaning of moral terms

39 Ibid.

or how moral remarks function. Is there a more effective way of overcoming this difficulty?

The third statement of ethical subjectivism is very similar to the second one, except that it emphasizes a distinction between fact and value. There is no 'moral order' out there. What we have is a physical order, an order of the sort of things that science can deal with. Now if there is no moral order out there, and if there is indeed a distinction between fact and value how can one coherently talk of an objective ethical theory? If there are no observable moral facts, how does one determine the validity of a normative ethical judgment?

Thus it seems that, whereas it is possible to dispose of some forms or varieties of ethical subjectivism, such as naive subjectivism and a simple version of the emotive theory, with relative ease, this is apparently not the case with some other forms of ethical subjectivism. The proper procedure for establishing which fundamental moral principles are true or false or justified or unjustified remains a serious challenge to ethical objectivism. To undermine all the varieties of ethical subjectivism, the ethical objectivist must be able to show that moral judgments are among the class of judgments whose truth is publicly determinable by any rational observer who is acquainted with the facts. He must be able to show how, in Kantian language, moral judgments can be valid for all rational beings.

It is evident from the foregoing discussion that Rawls cannot ignore some of the questions raised by ethical subjectivism in view of what he takes to be the nature of a theory of justice. By drawing an analogy between ethics and linguistic Rawls would have us believe that a normative ethical theory can be objective in the fashion of the social and empirical sciences. Thus he must be able to show that the truth of moral judgments can publicly be determined by any rational observer who is acquainted with the facts. Rawls' attempts to do this in all the three main arguments in A Theory of Justice. To what extent has he succeeded? This is one of the main questions to be investigated in this study.

PLAN AND METHOD OF APPROACH

This inquiry will be conducted in three parts. The first part focuses on the essential framework of A Theory of Justice. Here I consider Rawls on topics I find it central to consider in the light of some important remarks by his critics. My technique in this part is to bring out what I take to be in one way or another striking criticisms, some of which seem to me right minded, some importantly confused and some partly right and partly wrong. With those in view, I shall try to clarify fundamentally what is going on in Rawls' account and what should be said about these issues themselves. In doing this, I shall develop what in effect is

a dialogue between Rawls and his critics and then try to adjudicate the conflict and push the issue as far as I can.

Three themes will be considered in this part: the original position, Rawls' moral epistemology and the two principles of justice as fairness. In examining these themes, particular attention will be paid to the following questions: (1) Given its circumstances, especially the veil of ignorance and the motivational assumption of self-interestedness, can a rational choice of moral principles ever emerge from the original position? (2) Does Rawls' appeal to what he takes to be humanity's fixed and firm moral convictions about fairness implicitly commit him to a variant of relativism or subjectivism? and (3) Do the two principles of justice as fairness, by themselves, actually reveal our nature as moral persons? Why for instance is it impossible for an amoralist or a free-rider egoist to, in a certain sense, appeal to these same principles even though he does not have any regard for the moral point of view?

In part two of this study, I shall examine Rawls' theory of the good. My approach here consists in a critical probing of the various elements in this theory to see whether or not, Rawls is able, at least in this domain, to undermine some of the difficulties raised by moral scepticism. In other words, by articulating his account of the good, with special emphasis on such conceptions as the primary goods (including the good of

self-respect), rational life plans, principles of rational choice, the good of a person, the good of the sense of justice, and the congruence of goodness and justice, I want to determine whether Rawls' account is sufficiently determinate and unproblematic not to raise some of the central worries which are normally associated with moral scepticism. There are two general questions that are of special interest in this section. The first is about the notion of rationality. Has Rawls given us a generally satisfactory account of rationality, such that this concept may be seen to be plainly uncontroversial? And more importantly, can the socio-economic model of rationality to which Rawls appeals in his theory ever commit anyone to the moral point of view? The second question relates to the "Why be Moral?" question. Has Rawls' account of the good succeeded in showing that being a morally good person is always a good thing for anyone at all, such that everyone ought to desire to be a morally good person and not just a man of good morals?

In part three some proposals are put forward towards a revision of Rawls' account. More importantly, it will be argued that there is a need for a morally built-in notion of rationality within the framework of Rawls' theory. The basis for these proposals stems directly from the sort of thing Rawls is trying to achieve in his theory. In interpreting the initial situation for justice as fairness, Rawls wants

his characterization of his original position to capture the idea of a well-ordered society: a society in which individuals are to view themselves and one another as free and equal moral persons. To evolve such a society, the persons in the original position must, among other things, be presumed to be rational in the sense that they can understand and accept moral values as sufficient and final reasons for acting. As mutually disinterested persons, and being rational only in a morally neutral sense, the only reasons that Rawls' persons can possibly recognize for acting are prudential reasons of self-interest. In that case, it is very difficult if not impossible for them to understand or make sense of the motivation of an individual who acted out of concern for the welfare of others. Only a person with moral values can appreciate morality and moral reasoning.

Also, the very ideal of a well-ordered society that Rawls articulates in his account cannot be realized without a stronger and built-in conception of rationality. Rawls tells us towards the end of his book that the idea of a well-ordered society corresponding to justice as fairness "is itself a form of a social union. Indeed, it is a social union of social unions. Both characteristic features are present: the successful carrying out of just institutions is the shared final end of all the members of society, and these institutional forms are prized as good in themselves". Thus being a member of a

well-ordered society might require someone to sacrifice his personal interests in certain instances in order to promote the common final end of the society. But how can this be possible if Rawls' persons can only understand and accept reasons maximizing their own personal interests as the one and only basis for rational behavior.

On the whole, Rawls' theory of the good may have provided a plausible answer to the question "why should we be moral?" We should be moral because everyone, or at least most people can more effectively attain their human good when we are moral. But this question is not quite the same as "Why should I be moral?" To answer the latter, Rawls must be able to show that it is yet rational and intrinsically desirable for an individual to act from the moral point of view, even when he knows that he might suffer personal disadvantages by so doing. Rawls has not been able to do this nor is it certain that anyone can do it by merely appealing to empirical or non-moral conditions. There is no non-moral reason for being moral. Any attempt to defuse amoralism by pointing to the advantages of morality (whether these be collective or individual advantages) cannot be conclusive. A reflective and informed amoralist will be aware of these advantages as well as the need for a collective sense of morality. The point of his attack is the rationale for an unconditional individual allegiance to morality. Why should an individual who has

nothing to lose or gain act from the moral point of view? To meet this attack it must be shown that being moral is one of the necessary conditions of being a rational person. It is doubtful whether Rawls or anyone for that matter can prove such a proposition. At least it cannot be done without presupposing some ideal of a human person which is itself open to challenge.

The proposals suggested in this final section may indeed be useful and quite compatible with the sort of morality Rawls is trying to generate. However, the present framework of his theory cannot accommodate them. The original position would lose the deductive justificatory power Rawls had envisaged, if he were to introduce substantive moral notions into it. The derivation of Rawls' principles of justice will become trivial; and, the amoralist can even demand why he ought to be rational if rationality is taken in a moral sense. But it looks as if there is a kind of dilemma here. If one sticks to a morally neutral conception of rationality one cannot arrive at the principles of justice and if one enriches the conception of rationality one would seem to beg the question raised by the amoralist. In either case one does not defeat the amoralist. If this is the case, can we not simply ignore the fact of the amoralist or the subjectivist (trusting, as Hare does, that there would not be many of his type around) and proceed to do normative

ethics in a way that meets the considered judgments of the majority. While Rawls himself seems to have such a thought in mind (several passages in the final sections of his book indicate this) he has not taken it into account in deciding the approach of his work.

The inquiry will now begin with an examination of the notion of the original position.

PART I

THE ESSENTIAL FRAMEWORK OF A THEORY OF JUSTICE

CHAPTER ONE

THE ORIGINAL POSITION

PREVIEW: RAWLS' IDEA OF THE SOCIAL CONTRACT

Men being, as has been said, by nature all free, equal and independent, no one can be put out of this estate and subjected to the political power of another without his own consent. The only way whereby anyone divests himself of his natural liberty, and puts on the bonds of civil society, is by agreeing with other men to join and unite into a community for their comfortable, safe, and peaceable living one amongst another And thus that which begins and actually constitutes any political society is nothing but consent of any number of free men capable of a majority to unite and incorporate into such a society. And this is that, and that only, which did or could give beginning to any lawful government in the world.

John Locke, Second Treatise of Civil Government.

This passage articulates two of the fundamental beliefs of classical contractarianism: the belief that men are by nature free, rational and equal; and the belief that any form of association that they enter with their fellow men derives from their free consent and aims at protecting their individual liberties. Although there are substantive variations in the cluster of theories that make up this tradition, all the classical contractarians uphold these beliefs in one form or another. Hence the goal and purpose of the contract theorist, as Rousseau emphasizes, is to find out what form of association would appropriately "defend and protect with the whole force of the community the person and

property of every associate, and by means of coalescing with all, each may nevertheless obey himself, and remain as free as before"⁽¹⁾.

By situating justice as fairness in the tradition of the social contract, Rawls also shares the belief that men are by nature free, rational and equal; and that the principles and forms of social and political associations must derive from individual free consent. The conception of justice that most suitably defines a well-ordered society, he tells us, is one that would be unanimously agreed to in a hypothetical situation that is fair between individuals conceived as free, rational and equal moral persons⁽²⁾. However, while affirming these classical beliefs, Rawls' formulation of the social contract attempts two signal breaks with classical contractarianism. First, in trying as he does, to provide a moral basis for the ideals and practices of liberal democracy, Rawls, in effect, rejects the doctrine of neutrality held by his classical predecessors. And, secondly, while accepting individualism as a theoretical starting point of contractarian liberalism, he denies the predominant view that liberalism is necessarily wedded to the individualist conception of moral personality. Thus in his version of the social contract, Rawls tries to defend a communitarian ideal as against an individualist ideal of moral personality. It will become evident in the next few chapters just how far he has succeeded in this task. Here we focus on his attempt to

provide a moral foundation for contractarian liberalism.

What led Rawls to embark on such a project?

Rawls repudiates the attempt by some theorists to apply both the economic and the pluralist accounts of democracy to the actual constitutional process in the democratic society. Such application has grave limitations, he says, "insofar as political conduct is affected by men's sense of justice"; and, in any viable society, "just legislation is the primary social end"⁽³⁾. Hence, the basic structures of society cannot be morally neutral or inconsequential. These structures regulate the pattern of distribution of relative wealth, status and power. And, since the relative life prospects of any individual or groups depend on these basic structures, the structures themselves become the primary subject of justice and morality.

Taken together as one scheme, the major institutions define men's rights and duties and influence their life-prospects, what they can expect to be and how well they can hope to do. The basic structure is the primary subject of justice because its effects are so profound and present from the start. The intuitive notion here is that this structure contains various social positions and that men born into different positions have different expectations of life determined, in part, by the political system as well as by economic and social circumstances. In this way the institutions of society favor certain starting places over others. These are especially deep inequalities. (4)

1 Rousseau, Social Contract, bk 1, Chap.6.

2 Rawls, "Reply to Alexander & Musgrave" in Quarterly Journal of Economics, Vol. 88, 1974. p.637.

3 Rawls, A THEORY OF JUSTICE, pp.360-1, see also footnote p.360.

While Rawls' remarks seem obviously correct, one may yet wonder whether this is the way society ought to be. Why, for instance, should inequalities be allowed to exist at all in the society? Or, if somehow inequalities just come to exist, why should man not strive to eradicate them? Why should they be allowed to persist? In view of his commitment to certain ideals of liberalism, these questions may appear illegitimate or at least uninteresting to Rawls. The issue for him is not whether inequalities should exist or why they should exist. They are, as a matter of fact, part and parcel of society.⁽⁵⁾ Hence the question for him is : under what conditions can social inequalities be morally justified and permissible.

Thus, for Rawls, the fundamental problem of the social contract is to design a formula that yields a distributive structure whereby some patterns of unequal wealth, status and power can be seen to be morally justified. In addressing himself to this problem Rawls identifies two principles of justice which, according to him, would be acknowledged by equal and rational persons as legitimate and just criteria for designing and assessing basic institutions. Rawls' contractarianism

4 Ibid. p.7

5 But if Rawls accepts inequalities as a normal or inevitable condition of society, how is one to understand the doctrine of individual equality that echoes so strongly in his theory? Does it not sound like a paradox for a person to claim allain one breath that men are equal and that inequalities are the very status quo of society? This question will be closely examined when we come to review Rawls' principle of Equal Liberty.

(6)
from this standpoint is essentially methodological. Its aim is to show, under what conditions a society operating within the framework of liberalism (free enterprise, competitive markets, private property in the means of production, legal protection of freedom of thought and liberty of conscience, etc.) can be said to have a distributive structure that is morally justifiable. The central feature of this contractarianism is the original position. This is a hypothetical device employed by Rawls to identify and collate both the moral and morally relevant empirical premises from which the fundamental principles of morality are to be derived. Whether the original position is capable of fulfilling these roles is the subject of this chapter. Emphasis will be on two items: (1) the original position as a theory of moral analysis; and (2) the original position as a formula of moral justification. We begin with the former.

6 In this respect, Rawls' contractarianism is similar to that of Hobbes. Although Rawls' construction of the original position bears the semblance of an historical account, he does not intend it to be so interpreted. Unlike Rousseau and Locke, Rawls does not rest his theory on any examination of what the character of an actual pre-political condition or "state of nature" among men would be. He does not try to infer his principles of justice from what men, given their nature, would likely agree upon in the absence of an already existing government. Rather his contractarianism is only a thought experiment for restructuring the institutions of the liberal democratic society in order to bring them into conformity with the constraints of justice.

A THEORY OF ANALYSIS AND CRITICISM

Rawls freely constructs the character and circumstances of the parties to his purely hypothetical original position in order to lead to a conception of justice; that of justice as fairness. Among other things, we are to assume that every agent in this original position is rational; that is, each person seeks to maximize his well-being, though subject to the constraint that he is not envious of others. Also each person knows that he is subject to "the circumstances of justice".

(These are) the normal conditions under which human co-operation is both possible and necessary Hume's account of them is especially perspicuous For simplicity I often stress the condition of moderate scarcity (among the objective circumstances), and that of mutual disinterest of individuals taking no interest in one another's interests (among the subjective circumstances). Thus, one can say, in brief, that the circumstances of justice obtain whenever mutually disinterested persons put forward conflicting claims to the division of social advantages under conditions of moderate scarcity. (7)

While the agents are presumed to know the general facts about life situations, and human society in general, no one has any particular information about himself. They are to choose their principles of justice from behind a

"veil of ignorance" that ensures among other things that no one knows his place in society, his class, position or social status; nor does he know his abilities, his intelligence and strength, and the like. Nor, again does anyone know his conception of the good, the particulars of his rational plan of life, or even the

special features of his psychology such as his aversion to risk or liability to optimism or pessimism. (8)

By denying the members of the original position particular information about themselves, it becomes impossible for anyone of them to know who he is. And if nobody is able to identify himself, it follows that nobody would be able to identify others as well. No one knows what differentiates him from others. Hence each person can only see himself as anyone.

... it is clear that since the differences among the parties are unknown to them, and everyone is equally rational and similarly situated, each is convinced by the same arguments. Therefore, we 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. (9)

The principles that these parties will choose must in turn "match our considered convictions of justice or extend them in an acceptable way ... (so that) applying these principles would lead us to make the same judgments about the basic structure of society which we now make intuitively and in which we have the greatest confidence, or in cases where our present judgments are in doubt, would offer a solution which we can affirm on reflection." (10)

(8) Ibid. p. 137.

(9) Ibid. p. 139.

(10) Ibid. p. 19.

As it is for every other contractarian conception of justice, the interpretation of the initial situation is a very important factor in Rawls' theory. There are, he says, several possible interpretations of the initial situation and each interpretation not only leads to a different conception of justice, but to a different contract theory as well. In view of this fact, there are several different contract theories. And to assess the validity of any particular one, the first requirement, according to Rawls, is to show that there is at least one interpretation of its initial situation "which best expresses the conditions that are widely thought reasonable to impose on the choice of principles yet which, at the same time, leads to a conception that characterizes our considered judgments in reflective equilibrium"⁽¹¹⁾. Rawls' conjecture is that "for each traditional conception of justice there exists an interpretation of the initial situation in which its principles are the preferred solution"⁽¹²⁾. The most favoured or standard interpretation of the initial situation that leads to justice as fairness is the original position.

11 Ibid. p.121.

12 Ibid.

These two suppositions raise some "important questions concerning the objective validity of Rawls' theory. For any particular interpretation of the initial situation to be considered adequately justified, the conditions it expresses must be widely seen as reasonable. Also the principles emerging from it must be seen as the preferred solution under the circumstances and conditions articulated in the interpretation. But reasonable to whom and preferred by whom, and

Thus in its analytic role the original position offers a method for the comparative study of the notion of justice as fairness in relation to other conceptions of justice.⁽¹³⁾ By using such a device, we come to know the assumptions underlying a particular conception of justice by discovering a set of premises from which the principles of the conception can be derived. In this way we also come to see the degree to which our conception of justice depends on various formal requirements such as the conditions of generality, universality, finality, and publicity that apply to the principles. Also in this role the original position attempts to furnish a concrete model, as Norman Daniels points out, "for reducing a relatively complex problem, the moral choice of principles of justice, to a more manageable problem, the rational individual choice of principles". Now, to what extent can the original position deliver on these claims? How coherent is it as a theory of analysis? In discussing these questions we pay particular

on what basis? Are the conditions and circumstances of such a nature that any rational person regardless of his particular point of view would consider them reasonable? What would be the criteria of such reasonableness? These questions are particularly disturbing in view of some of the assumptions made by Rawls in the original position. If any random person can adequately choose and think for all, then the criteria of his choice and thinking must be unquestionably reasonable. Are there such criteria, and has Rawls been able to specify anyone of them? These questions will be examined closely in the next chapter when we come to examine Rawls' moral epistemology.

attention to the very possibility of isolating the most reasonable interpretation of the initial situation. Is such a task even theoretically possible?

ISOLATING THE MOST REASONABLE INTERPRETATION OF THE INITIAL
(15)
SITUATION

As seen earlier, Rawls supposes that "for each traditional conception of justice there exists an interpretation of the initial situation in which its principles are the preferred solution. (16) And in a different context, he asserts (17) that "different premises can yield the same conclusion". The latter assertion, we might suppose, would apply not only to ordinary logical arguments but to the principles of justice as well as the premises from which such principles are derived. Hence it would seem valid to conclude that for each conception

13 To illustrate this, Rawls demonstrates how the principle of average utility which directs society to maximize the average utility (per capita) can be derived. He postulates an initial situation in which the parties are single rational individuals behind a veil of ignorance, and who wish to maximize their own well-being. By waiving the problems of inter-personal comparisons of utility, and assuming that each person does not know who he is, (neither his natural ability nor his social position), the parties following the principle of insufficient reason in computing likelihoods, would opt to be governed by the principle of average utility. This seems to be the most preferred solution from this initial situation because it is rational for each hypothetical contractor to choose to enter a society in which his prospect is highest. If his reasoning is correct, such a society would be the one governed by the principle of average utility. (Rawls, Ibid, pp. 164-166). Rawls however rejects the principle of average utility on the basis of his interpretation of its initial situation. He calls it an "ethic of single rational individual (with no aversion to risk) who tries to maximize his own prospects". (Ibid. p.189).

of justice, it is possible to construct several different interpretations of an initial situation. David Keyt's example of an alternative derivation of the principle of average utility adequately supports this claim.⁽¹⁸⁾ In this example, Keyt appeals to a set of premises radically different from the one used by Rawls in the derivation of the same principle. Keyt's derivation begins as follows:

Suppose we postulate an initial situation in which the parties are perfect altruists who sympathetically identify with one another and with the possible descendants of one another. Each party chooses as if he himself will live through the experiences of all the persons with whom he identifies. Suppose, however, that the parties use as their principle of rational choice the intensity rather than the maximization principle. Thus they choose to maximize, not the total net balance of satisfaction in the totality of experience with which they identify, but rather the average net balance per unit of time. Suppose, finally, that the unit time they adopt is a lifetime, admittedly (like a calendar month) a unit of variable length. They calculate the average net balance of satisfaction in the series of lives with which they identify by dividing the total net balance of satisfaction in the entire series by the number of lives in the series. In this initial situation the parties, using the intensity principle as their principle of choice and a lifetime as their unit of measurement, will choose to maximize the average net balance of satisfaction per capita, which is to say that they will adopt as their principle of justice the principle of average utility. Thus we have an alternative derivation of the principle. (19)

14 Norman Daniels, Reading Rawls' op. cit., p.xix.

15 The critical comments in this section articulate some important criticisms made by David Keyt concerning the basic assumptions behind Rawls' appeal to the original position. "The Social Contract as an Analytic, Justificatory and Polemic Device", Canadian Journal of Phil., Vol. iv. No.2, 1974, pp. 244-252.

Apart from the modification in the principle of rational choice, the premises of Keyt's derivation are the same as the ones from which Rawls derives his principle of classical utilitarianism. ⁽²⁰⁾ Now if different interpretations of the initial situation can thus yield the same principle of justice, then we can assume, among other things, that it is theoretically possible, at least, to construct an interpretation of the initial situation which differs from Rawls' original position and yet obtain the same principles of justice and fairness.

While Rawls does not deal explicitly with this problem in his book, he is aware of its significance. If different interpretations of the initial situation could yield the

16 Rawls, op. cit., p.121.

17 Ibid. pp. 387-8.

18 Rawls' suggestion that his own way of deriving the principle of classical utilitarianism is not the only way of deriving the principle may also be seen as a support for this claim. (C.f. Rawls. p.26).

19 David Keyt, Op. cit., p.246.

20 To derive the principle of classical utilitarianism, a principle which directs society to maximize total utility, Rawls postulates an initial situation "which the parties are perfect altruists who sympathetically identify with one another. Each person views the pains and pleasures of others as if they were his own ... one chooses as if one will for certain live through the experiences of each individual seriatim as Lewis says, and then "sum up the result". (Rawls p. 189). All persons are conflated into one (p.191) and "the principle of rational choice for one man is taken as the principle of social choice as well". (p.189). While Rawls does not specify the extent of the sympathetic imagination of these perfect altruists, he warns that the classical principle becomes identical with the average principle if the size of the population remains constant.

same conception of justice, then some reason must be given why a particular interpretation is to be considered "most favoured". It is not just sufficient to choose one conception of justice simply because there is a reasonable interpretation of an initial situation that leads to it. Rather any one interpretation must be carefully scrutinized and assessed against other possible interpretations to see if it can emerge as the most reasonable interpretation. Rawls lays out how he means to handle this problem in a private letter to David Keyt.

He declares :

What we do is to associate with each conception the strongest argument for it (insofar as this same argument depends on an interpretation of the initial situation). This is fairly standard: a doctrine stands or falls according to the best reasons that can be given for it. The bad reasons we ignore. (21)

To present a real choice option for the parties therefore, they must sympathetically identify, not only with one another, but also with the possible descendants of one another. The principle of rational choice that Rawls uses in this derivation is one that directs a person to maximize "the net balance of satisfaction". (Rawls, p. 416). While this principle does not help generate the classical principle, there is no reason why the parties cannot adopt a different principle; for instance, a hedonist, as Keyt suggests, might decide to maximize the average net balance of satisfaction per year instead of total net balance over his entire life. Hence, the hedonist would choose a short life with a high average net balance of satisfaction over a long life with a larger total but a lower average. If he has to make a choice between a longer and a shorter life both of which contained the same net balance of satisfaction, he would choose the shorter, the more intense of the two. Keyt's position here seems quite supportable from a text in Aristotle's Nicomachean Ethics. Aristotle asserts that the good man "would prefer a short period of intense pleasure to a long one of mild enjoyment, a twelve month of noble life to many years of humdrum existence, and one great and noble action to many trivial ones". (E.N.11 69^a 22-25, Ross's translation.)

One of the initial difficulties that confronts the contract theorist, as it emerges from this brief remark, is that none of the premise sets which he might want to use to characterize his original position is logically equivalent to the set of principles he is interested in deriving. If he were able to associate with a particular conception of justice an interpretation of the initial situation which is logically equivalent to the principles of that conception, then he would simply ignore other sets of premises which entail the same principles. As Rawls does not envisage the possibility of such interpretations, he considers it sufficient that the theorist be able to rank the premises in his family of sets according to their reasonableness. Thus, if a particular set of premises is, or appears more reasonable than a second set to a group of persons, then it is ranked higher for them. When the premise sets in the entire family of sets must have been ranked, the one at the top becomes the most reasonable for this group of persons. For this group therefore, a single set of premises is associated with each conception of justice, and for them the conception is characterized by this premise set and stands or falls with it. The most obvious question at this point is the criteria of reasonableness. What considerations must the group appeal to in ranking their premises? Would such considerations, if any, make equal appeal to any random group of persons?

Rawls indicates some of the factors that should be considered in such a ranking exercise. He says:

... a premise set should be : consistent, including only true (or approximately true) general beliefs, enumerate plausible formal conditions on principles, specify sensible motivations, etc. (22).

There are, no doubt, certain positive things to be said for some of these constraints. For instance, in assessing the validity or soundness of any argument at all, one of the things anyone would look for is the consistency of its premises and conclusion. There is hardly any rational ground or circumstance that would justify the acceptance of an argument that is known to be inconsistent. Also, an argument which is known to include untrue general beliefs, if indeed there are such things as general beliefs, can hardly be taken seriously. However, while the rationale for these two constraints seem obvious enough, it is still doubtful whether the criteria suggested by Rawls can lead to a consensus among rational persons as to which particular interpretation of the initial situation is most reasonable. Rawls does not indicate what criteria must be used in order to determine whether or not a particular motivational assumption is sensible. Even if we are to assume that the group of persons he has in mind refers to the set of all rational persons, and, that nothing would appear reasonable to a rational person except it is indeed reasonable, we still cannot anticipate a consensus. The judgments of rational persons are not totally

free of individual bias or idiosyncrasy. Keyt is plainly correct in noting that, whether a particular formal condition is more plausible than an alternative one, or one motivational assumption more sensible than another would depend largely on individual intuition. And rational persons can reasonably disagree with one another on such issues, especially when these issues relate directly to psychological assumptions. Keyt draws a useful analogy between what Rawls tries to do here and what some logicians attempt to do in respect to axioms. He says :

Finding the most reasonable set of premises from which the principles of a given conception of justice can be derived is similar to finding the most elegant set of axioms to be both complete and sound, to yield all and only tautologies as theorems. There are many sets of axioms which satisfy these criteria, some of which are more elegant than others. The set that Frege proposed in the Begiffsschrift, the first complete and sound set to be discovered, is indeed elegant, like all of Frege's work; but still more elegant axiomatizations have subsequently been discovered. It would be rash to suppose that the premise sets that Rawls has associated with the versions of utilitarianism are in a better position with respect to reasonableness than Frege's axioms with respect to elegance. We have not yet had a chance to examine other derivations and to compare them with his. (23).

Keyt's remarks are well taken; and the original position does seem to lose considerable force as a theory of analysis and criticism. For the theory to succeed as Rawls conceives

23 Ibid.

It is important to note that Keyt's remarks here tend to support Nowell-Smith's charge of platonism in Rawls. However, it might still be necessary to take a close look at Rawls' notion of reflective equilibrium in order to determine whether this is actually the most plausible way of interpreting his theory. This question will be examined in chapter two.

it, it must be possible to show that any random individual at all, provided he is rational, would be able to identify the most reasonable interpretation of the initial situation from all the possible interpretations that could lead to a particular conception of justice. The very prospect of such a project requires three basic assumptions which are not easy to validate. As Keyt points out, it must be shown that:

1. For every conception of justice, there is at least one interpretation of the initial situation that is maximally reasonable;
2. For every conception of justice there is at most one maximally reasonable interpretation of the initial situation; and
3. Within the family of premise sets that yield a conception of justice, the maximally reasonable set can be isolated and recognized. (24)

The rationale for these assumptions is not too difficult to see. In respect to the first, let us suppose that the members of the family of premise sets that characterize a conception of justice approach a limit of reasonableness without ever reaching it. Such a supposition is quite legitimate insofar as no member of this family of sets is logically equivalent to the principles of the given conception of justice. Obviously, the most reasonable premise set in any family of sets would be the set that is logically equivalent to the principles of the conception itself. However, in the absence of such a premise set, a situation is conceivable in which no premise

24 Ibid. p.250.

would stand on top of the rank. We can always anticipate a further premise set that is more reasonable than any given members of the family. Thus, should one reject a conception of justice because some unreasonable premise set occurs somewhere along the line of its derivation, as long as this premise set is not logically equivalent to the principles of the conception, one would always face the thorny problem of anticipating another premise set from which the same principles can be derived which lacks this unreasonable element. In such a frame of mind, it becomes difficult, if not impossible, to reject a particular conception of justice simply because there exists some unreasonable element in the family of sets that yields the conception. For Keyt to have been able to derive the principle of average utility at all, without appealing to the principle of insufficient reason, seems to prove this point. Although parts of his derivation might be objectionable, the attempt itself does show the possibility of appealing to different premises and yet arriving at the same principle. Contrary to Rawls' claim therefore, it is not a sufficient condition for discarding the principles of a conception of justice just because one discovers certain unreasonable element in the initial situation leading to it. We can always entertain the possibility of appealing to a new set of premises that might yield the same principles. Such a possibility can only be ruled out on the assumption that there is at least one interpretation of the initial situation that is maximally reasonable. But this assumption is by no means easy to demonstrate.

Granted that this assumption can be validated and it is indeed correct, the second assumption mentioned above raises another serious problem. As noted above, the successful use of the original position in an analytic role further requires an assumption that there is at most one maximally reasonable interpretation of the initial situation for every given conception of justice. The rationale for this condition seems straightforward. For instance, suppose that a given conception of justice is derivable from two different sets of premises each of which is considered maximally reasonable. In such a situation, it is impossible to characterize the conception with any of the two sets of premises. As we saw earlier, Rawls, on the basis of his derivation, characterizes the principle of average utility as "the ethic of a single rational individual (with no aversion to risk) who tries to maximize his own prospects...."⁽²⁵⁾ But from Keyt's alternative derivation, the principle can also become the ethic of a perfect altruist who follows an intensity principle of rational choice. Now suppose both of these derivations are equally reasonable, then none of them would qualify as the characterization of the principle of average utility. Hence, Rawls cannot claim to have produced the characterization of the principle of average utility anymore than Keyt.⁽²⁶⁾

25 Rawls, A Theory of Justice, p.189.

26 As Keyt indicates a problem arises here because "the relation of being more reasonable than does not simply or (linearly) order the family of sets associated with a given con-

The third assumption listed above is equally crucial. The success of the original position as a critical device further requires an additional assumption that within the family of premise sets that yields the principles of a given conception of justice, the maximally reasonable set can be isolated and recognized. While the rationale of this constraint is rather obvious, it is not clear how it could or whether it could even be met at all. It requires that the theorist be familiar with all the members of the family or at least know that any member that he is unacquainted with is no more than equally reasonable if compared to some member that he is familiar with. The problem that the theorist faces here does not seem easily resolvable. It is not even clear how he can begin to make a move.

If the issues raised so far are valid, Keyt's criticism of the original position would seem to stand. In particular, the validity of the assumptions behind its analytic use is at best doubtful. These assumptions are crucial, and, their truth must be established beyond question. Of course, this might appear too strong a Cartesian demand. A Peircean fallibilist might respond to it by saying that almost nothing gets

ception of justice. Two distinct members of such a family may be equally reasonable and thus stand at the same rank. (Which of the two premise sets that have been proposed for the principle of average utility is more reasonable?) The relation of being more reasonable than only partially orders the family of sets associated with a given conception of justice. It operates here like the relation of having more members than among sets rather than like the relation of being larger than among natural numbers. Distinct sets can be the same size but not natural numbers". (Keyt. op. cit., p.249.)

established beyond question; but that if these three assumptions can be shown to be more plausible and reasonable than their alternatives in some non-question begging way that would be sufficient. Such a response does have some substance, and, perhaps we need not go on a Cartesian quest for certainty. We might say that, as an analytic device, it suffices that the original position enables us to understand some of the important premises that lead to a particular conception of justice. And by comparing this set of premises with other sets of premises that could lead to other conceptions, we could make a reasonably secure judgment in favour of one conception over others. In such circumstances the question of isolating the most reasonable interpretation of the initial situation no longer presents insurmountable difficulties.

An approach such as this does seem plausible, and would seem to restore some force to the original position. However, it must be pursued with caution in view of some of the fundamental assumptions of justice as fairness. Rawls' intent in constructing the theory is not merely to offer a viable or plausible conception of justice. Rather it is to present "an Archimedean point for judging the basic structure of society", according to the principles of social justice. Rawls is convinced that these principles can both be specified and rendered determinate through a process of rational inquiry. He believes that a rigorous thought experiment which appeals only to our knowledge of the general facts about human society can lead to

the discovery of a single, structured set of principles which underlie our sense of justice. Also, with the veil of ignorance covering up all the particular contingencies that set limits on the rational capability of individuals, each member of the original position is equally rational and similarly situated such that each is capable of being convinced by the same argument.

Now in face of strong assumptions such as these, there is hardly much room for one to manoeuvre in the direction of Peircean fallibilism. While a Peircean fallibilist may still want to discover a point from which he could reasonably assess the systems of moralities, he would not consider such a point as the Archimedian point. And if the choice and thinking of one person selected at random would coincide with the choice and thinking of everyone else, then such a one must be capable of knowing what everyone could possibly know about any given situation. Again such an assumption seems too strong for one to make in the context of fallibilist thinking. Thus one may conclude that even though the original position does indicate a useful procedure for analyzing and criticizing the various conceptions of justice, the extent in which Rawls wants to use it as a critical device does raise some serious difficulties. Granted however, that there is not much anyone could do to remove such difficulties, it would still be premature to discard the thought experiment of the original position outright. As mentioned earlier, Rawls wants the original

position to play an analytic as well as a justificatory role. In actual fact, Rawls' emphasis is on the latter role. Therefore, until this aspect has been carefully scrutinized, an assessment of the success of the entire device would be rather inconclusive. Thus in what follows we take a critical look at the original position in its justificatory role. The discussion here will focus on some of the important issues raised by several of Rawls' critics.

SECTION II

A THEORY OF MORAL JUSTIFICATION

In formulating the conditions and circumstances of the original position, Rawls wants the choice of the set of principles emerging from it to constitute a justification for these principles. Hence for him, the principles of justice as fairness are to be considered justified if they could be shown to be the principles that would be selected by any and all rational persons in an initial situation of equality and freedom, in which the basic structure of society is to be designed from scratch.⁽²⁷⁾ The basic condition that must be fulfilled in the choice of these principles is that both the premises from which they derive and the consequences that derive from them must match the considered judgments of rational persons as being reasonable and fair. Rational persons must, on due reflection, acknowledge the principles as the ones with which they wish to adjudicate their competing claims in a well-ordered society. Thus in its justificatory role the original

position presents a model for procedural fairness. It advances a set of conditions and constraints which would appeal to every rational person on due reflection as both fair and just. Now to what extent does it succeed in this role? This is the issue to be determined for the remainder of this chapter. Emphasis would be placed on two questions which have been raised by several critics of Rawls: (1) Is the relation between the original position and the two principles of justice as fairness truly a deductive one? ; and (2) Even if the parties in the original position do in fact choose these two principles, does the fact of such a choice by itself articulate the moral virtue of fairness? The latter question is rather crucial in view of Rawls' rejection of the neutrality doctrine and his subsequent attempt to provide a moral foundation for contractarian liberalism. As was mentioned earlier, Rawls does not view the social contract in exactly the same way as Locke or Rousseau. The form of association he is interested in, is not one that merely protects our individual rights and liberties. Rather he wants a social contract that appropriately defines the notion of a well-ordered society: a form of association that truly reflects our nature as moral persons. Our nature as moral person is revealed not by our

27 The term justification in this context does not connote mere logical proof. Mere proof, Rawls says, "is not a justification. A proof simply displays logical relations between propositions". (Rawls, p.581).

individual conceptions of the good, but by the principles we choose to govern our life pursuits. Thus the crucial issue to be determined here is whether the choice of the principles of justice as fairness can even emerge from the original position; and, if it does, does it portray the parties as moral persons? It might be useful to make a further clarification on the features in the original position before proceeding directly to these questions.

(28)

From the three statements cited earlier, the features of the original position can be broadly divided into two kinds: constraints on knowledge and constraints on motivation. The constraints on knowledge, summarily stated, are that the hypothetical contractors do not know their place in society, their individual talents and bents, or their "conception of the good", which includes such items as their particular sources of pleasure, their life ambitions, and their religion or other beliefs. Also, in order to deal with those problems which may arise as a result of the "just savings rate", the contractors are denied the knowledge of what stage of economic development their society has reached. Rawls refers to these constraints as a "veil of ignorance". In regard to the motivational assumptions, the contractors in the original position are said to be rational although they are not altruistic. They are mutually disinterested in one another: each person merely concerns himself

to further his own "conception of the good" though he does not --because of the constraints on knowledge--know what content this conception will have. Because of this limitation, the contractors are denied any substantive sense of justice even though their conception of the good might eventually include a concern for the welfare of certain persons for whom they feel affection or special responsibility. Thus, as Brian Barry observes, "a man cannot, in the original position, take as his end the idea that everyone's welfare should be increased as much as possible, or say, that he would like a certain distribution of goods or utilities for its own sake".⁽²⁹⁾

From these conditions of the original position, Rawls attempts to deduce his two principles of justice as fairness. It is important to note here what precisely Rawls claims to be deducible. For instance, it would be inaccurate to understand him as attempting to provide a deductive proof of the proposition that whatever principles would be chosen in a suitably characterized original position are principles of justice. As Barry rightly observes, this "part of the argument depends on the plausibility of the idea that "the conditions of justice" are conditions of partial co-operation and partial conflict and the assertion that the conception of an 'original position' captures these features of the 'conditions of justice' while excluding irrelevancies"⁽³⁰⁾. Thus Rawls' claim to have established

29 Brian Barry. op. cit., p.11.

30 Ibid.

a deductive relationship only applies to one link in the chain: we are to proceed from an initial situation interpreted in one particular way (that is, as Rawls' original position) to the two principles of justice as fairness in their lexical order. Thus the statement of Rawls' deduction may be put summarily as follows: if the principles of justice are those principles which rational and mutually disinterested persons would choose in an initial situation of equality and freedom and under a veil of ignorance then, the two principles of justice as fairness are the principles of justice. ⁽³¹⁾ The goal of this deduction, as Rawls asserts, is to "strive for a kind of moral geometry with all the rigor ⁽³²⁾ which this connotes". However, as with the axioms of geometry or even logic, the postulated conditions of the original position have to be accepted first. The axioms of a system cannot themselves be proved within the system. Like metaphysical pos-

31 To anticipate, the central aim of the argument in this chapter is to show that the statement of Rawls' deduction is essentially inadequate. To the extent that Rawls himself takes justice as involving the concept of fairness, and fairness as a moral concept, then the principles of justice are not those principles which rational and mutually disinterested persons would choose. Rational and mutually dis-interested persons (self-interested persons) do not have the capacity to act from the moral point of view, hence they cannot choose the principles of justice as fairness as moral principles.

32 Rawls, op. cit., p.121.

Perhaps it should be pointed out here too that even though Rawls wants his arguments to be strictly deductive, he does not pretend to have achieved this goal in "A Theory of Justice". What he does usually is to present schemes for deductive arguments and not those hoped-for deductions themselves.

tulates, they are only criticizable in relation to the overall success of the system. Thus, our very first task here is to see whether the conditions of the original position do actually lead to the two principles of justice as fairness. In other words, is it possible to construct an interpretation of the original position such that the two principles of justice as fairness may be seen to derive from it? We examine this question with particular reference to Henry Shue's analysis of the logical structure of justice as fairness in his paper "Justice, Rationality and Desire".⁽³³⁾

Thus it would seem inappropriate to raise the question whether or not, Rawls has been successful at every stage of his projected deduction of the principles of justice as fairness. He admits the limitations of his arguments quite explicitly: "The argument aims eventually to be strictly deductive ... We should strive for a kind of moral geometry with all the rigor which this name connotes. Unhappily, the reasoning I shall give falls short of this, since it is highly intuitive throughout". (Ibid). This statement of confession by Rawls would seem to make it rather difficult to sustain some aspects of Nowell-Smith's criticism of the original position.

³³ Henry Shue, "Justice, Rationality and Desire", The Southern Journal of Philosophy, 13, 1975-6.

THE DERIVATION OF THE TWO PRINCIPLES OF JUSTICE AS FAIRNESS.

One distinctive structural feature of A Theory of Justice is the reverse logical order in which Rawls presents his theory. The exposition of the theory begins with an account of its conclusions (the two principles of justice). It then proceeds in a reverse deductive order to the premises that support these conclusions. Thus it emerges that early in the book, one sees the two principles of justice as fairness fully articulated. But when the reader has come to the third and final part of the book, only then does he begin to confront the deductive arguments for these principles. An illustration of this is the equal liberty principle. As Henry Shue notes, this principle is the concluding step of a projected deduction containing three major logical moves from its initial premises. He specifies the most important proposition at each stage of the deduction as follows :-

Principles of motivation → Principles of Rational choice → Primary Goods → Principles of Justice (36)

OR

Aristotelian principle → Principle of Inclusiveness → Self Respect → Equal liberty Principle

34 This approach might raise some suspicion of circularity. It is almost a consistent pattern all through Rawls' argument that he presents his conclusions first before going to display the premises for them. There are instances in which one has the feeling of being compelled to accept his premises not because they appear reasonable in themselves but because they imply the conclusion he has earlier stated. However,

Perhaps the ordering of these propositions might raise some questions. In fact what Shue takes to be the third inference in the diagram -- the inference from the primary goods to the principles of justice -- is seen by several of Rawls' critics as the basic inference of his argument.⁽³⁷⁾ It does not appear that the latter view is very correct. While the assumption about primary goods is indeed a crucial premise in the derivation of the two principles of justice as fairness, there is ample evidence to support Shue's contention that this is not the logical starting point or the most basic inference of Rawls' derivation. This claim can be supported by the theoretical explanation given by Rawls for the list of the primary goods. He says:

The thin theory of the good used in arguing for the principles of justice is restricted to the bare essentials. This account of the good, I call the thin theory; its purpose is to secure the premises about primary goods required to arrive at the principles of justice. (38)

despite this unusual logical procedure, Rawls' arguments seem yet immuned to the charge of circularity. There is no doubt that Rawls' interest is in the two principles of justice, and these happen to be the conclusions of his argument. One may also assert that it is because Rawls is first interested in these conclusions that he subsequently becomes interested in these premises. But it should be noted that the move from these conclusions to their premises is based not on a logical interest but on a psychological interest. As Henry Shue rightly argues, while Rawls' "interest in the premises which may lead him to use these premises in his argument derives psychologically from his interest in the conclusion the interest itself depends on a belief in an independent logical connection by which the conclusion may be deducted from the premises. Thus, the "because" in 'the premises are acceptable because they imply the conclusion' is a psychological, not a logical 'because'". (Shue, op. cit., pp.90-1).

35 Ibid, p.89.

36 Ibid,

It seems rather evident from these remarks that, while there are considerations about primary goods which logically support the two principles of justice, there is also a "thin theory of the good" that supports the premises about the primary goods. The thin theory must be seen as an essential element in the deduction in order to forestall the criticism of circularity. Rawls wants to argue that only an equal distribution of the primary goods could be just and fair. Before he can even begin to support this proposition, he must show why the primary goods are so crucial: why they are things that every rational person must want if he wants anything at all. In other words, Rawls must first show that the things he refers to are indeed "primary goods" before he can proceed to argue for their equal distribution. Rawls seems to anticipate the need for this when he says :

The primary goods can be accounted for by the thin theory of the good.... that liberty and opportunity, income and wealth, and above all self-respect are primary goods must indeed be explained by the thin theory. The constraints of the principles of justice cannot be used to draw up the list of primary goods that serves as part of the description of the initial situation. The reason is, of course, that this list is one of the premises from which the choice of the principles of right is derived. To cite these principles in explaining the list would be a circular argument. (39)

37 Both Michael Teitelman and Thomas Nagel hold this view. c.f. Thomas Nagel, "Rawls on Justice" in Reading Rawls,

38 Rawls, op. cit., p.396.

39 Ibid. p.433-4

Therefore, it appears that there are certain premises in Rawls' overall deduction which are logically prior to the premise about the primary goods. The logical starting point of his deduction is the thin theory. The assumptions in the thin theory generate the index of the primary goods, while the index of the primary goods leads ultimately to the two principles of justice as fairness. The main features of the thin theory are the psychological principles of motivation which Rawls refers to as the Aristotelian principle together with the three principles of rational choice. A detailed discussion of these features will not be undertaken here. The thin theory itself is one of the central elements to be examined in part two of this study. Hence the task in this chapter is simply to characterize the type of reasoning which underlies the original position. The aim is to specify at least one interpretation that could make the derivation of Rawls' principles logically compelling. In the course of my exposition, an attempt will be made to indicate some of the main difficulties implicit in the basic assumptions in Rawls' argument.

The first inference in Rawls' deduction derives from a generalization about a higher-order desire among human beings. Rawls designates this higher-order desire as the Aristotelian principle. The first statement of the Aristotelian principle occurs in the discussion of one of Rawls' three principles of rational choice: the principle of inclusiveness. It states that:

One long-term plan is better than another for any given period (or number of periods) if it allows for the encouragement and satisfaction of all the aims and interests of the other plans and for the encouragement and satisfaction of some further aim or interest in addition. The more inclusive plan, if there is one, is to the preferred. (40)

The statement of the Aristotelian principle itself is contained in one of the two considerations Rawls advances in support of the principle of inclusiveness in the long-term case. Here he maintains that one of the principles of rational choice is to be adopted because it follows from a general assumption about what human beings want, and that the case of the other rational principles are analogous. He says:

The other consideration is that, in accordance with the Aristotelian principle... I assume that human beings have a higher-order desire to follow the principle of inclusiveness. They prefer the more comprehensive long-term plan because its execution presumably involves a more complex combination of abilities. The Aristotelian principle states that other things equal human beings enjoy the exercise of their realized capacities (their innate or trained abilities), and that this enjoyment increases the more the capacity is realized, or the greater its complexity. A person takes pleasure in doing something as he becomes more proficient at it, and of two activities he performs equally well, he prefers the one that calls upon the greater number of more subtle and intricate discrimination. Thus the desire to carry out the larger pattern of ends which brings into play the more finely developed talents is an aspect of the Aristotelian principle. And this desire, along with the higher-order desires to act upon other principles of rational choice, is one of the regulative ends that moves us to engage in rational deliberation and to follow its outcome. (41)

40 Ibid, p.413.

41 Ibid, p.414.

One important question that one ought to keep in mind here is whether wants themselves can or should be subject to rational criticism. It does not seem that we can rightly avoid such

The logical dependence of the principle of inclusiveness on the Aristotelian principle is well articulated in another passage when Rawls asserts that it is "evident that the Aristotelian principle contains a variant of the principle of inclusiveness".⁽⁴²⁾ The Aristotelian principle is thus seen by Rawls as literally "containing" the principle of inclusiveness. Hence one correct way of stating the former is that "human beings have a higher-order desire to follow the principle of inclusiveness."⁽⁴³⁾ On the basis of the Aristotelian principle, therefore, it seems plausible to specify one of the objects of human desire as a desire to follow the principle of inclusiveness. It seems possible therefore to derive the principle of inclusiveness from the statement of the Aristotelian principle.⁽⁴⁴⁾

Of course the question yet remains whether the Aristotelian principle itself is well-founded. For instance, Brian Barry attacks it on the grounds that neither of the two possible interpretations that could be made of it would suit Rawls' purpose. If the principle is taken as partly defining

considerations in assessing Rawls' account. Rawls himself tries to deal with the question in his statement of the principle of "deliberative rationality". Of all the life plans that are consistent with the counting principles as well as other principles of rational choice, "one is to choose that life plan which one would choose as the plan one would most want, after careful reflection, in the light of a full knowledge of the facts and a vivid awareness of the facts, as that life plan which would best realize one's most fundamental desires". (Rawls, p.417). How sufficient a basis this statement is, will be determined when we come to examine the thin theory.

the notion of rationality, then Rawls would inevitably end up with some variant of perfectionism; a position he openly repudiates. On the other hand, should the principle be read as an empirical generalization then, as Barry contends, it could be shown to be plainly false. The general trend in the North American society would disconfirm it, if it is taken as an empirical generalization.

Barry's worry seems to be well-founded. 'If the Aristotelian principle plays such a crucial role then its status or basis of acceptance must be clearly defined.' However, it does not appear that the two positions he has indicated are the only options open to Rawls. The principles could be read as an empirical tendency statement. In which case, it need not be true of every particular individual before it can function as a valid premise in Rawls' deduction. This of course still leaves the problem of relevance. If it is a mere tendency, why should it play such a crucial role in the construction of a theory of justice? Why is it any more important than other human tendencies? These questions will be carefully examined when we come to re-examine the Aristotelian principle in the context of Rawls' thin theory of the good.

42 Rawls, p.427.

43 Ibid, p.414.

44 Perhaps it is necessary to indicate once more what Rawls is claiming in this deduction. What the statement of the Aristotelian principle does imply is that human beings tend to act as if the more inclusive plan is to be preferred. This does not imply that the more inclusive plan ought to be preferred. What Rawls has done here, as Henry Shue points

For the purpose of the discussion in this chapter, it suffices to note what role the principle plays in the derivation of Rawls' principles. By assuming that human beings have this higher-order desire, Rawls is able to infer the rational principle of inclusiveness.

Our next task is to examine the inference from Rawls' principles of rational choice to the assumption about the primary goods. Now, granting the correctness of the Aristotelian principle -- that people act as if the more inclusive plan is to be preferred -- how does this fact imply the statement of the primary goods? It might be useful here to be more specific in respect to the primary goods. How for instance, could it be shown that the principles of inclusiveness implies the primary good of self-respect? Why would a person who naturally seeks to realize his innermost and more complex abilities necessarily commit himself to seeking the primary good of self-respect?

out, might be termed as a deductive explanation rather than a deductive justification. It is an attempt to explain the rational principle of inclusiveness by deducing it from a theoretically more basic proposition. The essence of Rawls' claim is that given the Aristotelian principle people will in fact follow the rational principle of inclusiveness. Now, while the inference itself may appear quite legitimate if construed as an explanation, one may yet worry how it fits into Rawls' overall program of producing an Archimedian point for judging the whole of morality. It is important to know what people will likely do given certain circumstances or traits of their nature. But for normative ethics, the crucial thing is not what people will do but what they ought to do. Of course Henry Shue would like to play down this problem by insisting that even though Rawls' entire program is committed to a deductive justification, this does not mean that every step in his argument must aim at a

The way in which Rawls conceives the notion of self-respect would seem to offer a clue to this projected deduction:

We may define self-respect (or self-esteem) as having two aspects. First of all, ... it includes a person's sense of his own value, his secure conviction that his conception of his good, his plan of life, is worth carrying out. And secondly, self-respect implies a confidence in one's ability, so far as it is within one's power, to fulfill one's intentions. When we feel that our plans are of little value, we cannot pursue them with pleasure or take delight in their execution. Nor plagued by failure and self-doubt can we continue in our endeavours. (45)

Thus, Rawls postulates a necessary relationship between self-respect and the Aristotelian principle:

I assume then that someone's plan of life will lack a certain attraction for him if it fails to call upon his natural capacities in an interesting fashion. When activities fail to satisfy the Aristotelian principle, they are likely to seem dull and flat of competence or a sense that they are worth doing. A person tends to be more confident of his value when his abilities are both fully realized and organized in ways of suitable complexity and refinement. (46)

deductive justification. Perhaps this is so. But judging from the crucial role the Aristotelian principle is supposed to play in the entire deduction, one wonders which other premise could generate an ought if the Aristotelian principle is not supposed to. And, with such a deep psychological assumption in what sense is Rawls' theory normative?

45 Rawls, op. cit., p.442.

Rawls' conception of self-respect in this passage raises some worry. It is not so evident that the notion of self-respect is exactly the same thing as the notion of self-esteem. While both concepts do indeed imply confidence and satisfaction in one-self, self-esteem refers mainly to one's good opinion of oneself: projecting one's ego as a distinctive type of person. The notion of self-respect, on the other hand, refers mainly to confidence in one's own worth as a human being and a concern to maintain it. Hence to cause or make someone lose his self-respect is to make him feel diminished as a person. A loss of self-esteem is not to feel diminished as a person, but to feel diminished as the particular type of person one projects

The inference from the Aristotelian principle (and the principle of inclusiveness) to the primary good of self-respect is rather obvious from these two passages. As characterized by Rawls, the notion of self-respect becomes a necessary condition both for the Aristotelian principle as well as the principle of inclusiveness. Before human beings can come to have a higher-order desire to follow the principle of inclusiveness, it is necessary that they desire the primary good of self-respect. In which case, self-respect is a necessary condition for committing oneself to inclusiveness. Hence the principle of inclusiveness can be said to imply the primary good of self-respect.

Again, it is important to note here that the deductive relationship goes only in one direction : from the principle of inclusiveness to the primary good of self-respect, and not vice versa. There might be persons who desire the good of self-respect but who do not have any concern for the principle of inclusiveness or the higher-order desire stated by the Aristotelian principle. Here we might think of Rawls' hypothetical case of a mathematical genius who finds his greatest fulfillment not in solving mathematical problems but in counting the geo-

and conceives of one's self. Thus what Rawls has indicated here would more appropriately define the notion of self-esteem and not self-respect. As we will come to see later, while self-respect may indeed be conceived as a primary good by all rational persons, we cannot say the same for self-esteem. The latter might easily be taken for vanity.

metrical shapes of blades of grass in well-trimmed lawns. Such a person may still desire the good of self-respect even though he does not have any inclination toward the realization of his innermost and most complex abilities. Thus the point of Rawls' deduction is that the desire to follow the principle of inclusiveness necessarily implies a desire to affirm one's own sense of self-respect. His belief is that in seeking to realize our innate and complex abilities, we are at the same time seeking to assert our sense of self-respect. Therefore, the assumption about the primary good of self-respect can be supported by appealing to the principle of inclusiveness as contained in the Aristotelian principle.

The inference in this argument does appear compelling at least from a logical standpoint. If it is indeed the case that the desire for self-respect is a necessary condition for the desire to follow the principle of inclusiveness, then the latter obviously implies the former. ⁽⁴⁷⁾ Nevertheless, the basic assumption in this argument is still open to challenge. For instance, why must we accept the desire for self-respect as the necessary condition for the desire to follow the principle of inclusiveness, and not the other way round? Is it so plainly untrue that someone could prefer and actively seek more complex and demanding tasks without a concern for self-respect? Why must it be the case that anyone who is committed to inclusiveness (reasons from that principle) must also have self-respect?

⁴⁷ The argument is of this form: If Q is a necessary condition for P, then P implies Q.

It does not seem very helpful merely to refer to the way Rawls defines the notion of self-respect at this point. Such an empirical claim cannot be settled by mere definition. It has to be put to an empirical test. And the result of such a test might very well go against Rawls.

However, there is a more fundamental difficulty with this derivation. The latter is about the relevance of self-respect vis-a-vis life plans. If it is true, as Rawls' example cited two paragraphs above seems to suggest, that a person can have a desire for self-respect without any inclination towards the Aristotelian principle or the principle of inclusiveness, how does one represent the notion of self-respect in such a situation? For instance, in what does the self-respect of a clover leaf picker consist? Of course, if we take the notion of self-respect in the sense of the value a person attaches to himself because he is a person, we might yet be able to say in some vague sense that a person can still desire self-respect, even if he does not desire to accomplish anything apparently useful or challenging. But Rawls' definition of the concept indicates a tight link between a person's life plan and the way he conceives of himself. The secure conviction that a person's conception of his good (his plan of life) is worth carrying out is the very basis of his sense of his own value. Admittedly, the clover leaf picker could also think that his plan of life is worth carrying out. Hence, he might also desire the supreme primary good of self-respect in the original position in order

to be able to accomplish this plan in real life. But if this is indeed the case, if self-respect is just as instrumentally important for becoming a clover leaf picker as it is for becoming a nuclear physicist, then this fact alone ought to tell against it to the extent that one is engaged in normative discourse. This question will be re-examined in some detail in chapters three and four. Meanwhile, it would seem evident from the questions raised in the preceding paragraphs that, while the inference from the principle of inclusiveness to the primary good of self-respect does appear to be logically compelling, the basic assumption in this deduction still needs to be argued for. Such an argument ought to take into account how the notion of self-respect is generally manifested by people in real life and not just how Rawls wants to define the term. We now turn to the final step of the deduction, that is, the inference from the primary goods to the principles of justice.

The final argument, in the derivation of Rawls' principles of justice centers on the assumption that there is a certain class of goods (primary goods) that every rational person ought to desire. The most important primary good is the good of self-respect. Hence for anyone to be rational and to desire anything at all, one must desire the primary good of self-respect. But to guarantee equality of self-respect, basic liberties must rank prior among other primary goods, and must be equally distributed.⁽⁴⁸⁾

48 The priority of equal liberty principle will be critically examined in chapter III when we come to probe Rawls' two principles of justice. Our primary concern at this point

The account of self-respect as perhaps the main primary good has stressed the great significance of how we think others value us ... The basis for self-esteem in a just society is not then one's income share, but the publicly affirmed distribution of fundamental rights and liberties. (49)

This brief remark indicates the basic thrust of Rawls' argument. The argument articulates a connection between the basic liberties and the main primary good of self-respect. Its general form may be characterized from two standpoints. From one standpoint, it takes the form of maintaining that a guarantee of the priority of liberty is the most effective means available for guaranteeing self-respect. Since the members of the original position are presumed to be rational (they are capable of taking the most effective means toward the realization of their desired end), they would give priority to basic liberties over all other primary goods once they have given the first priority to self-respect. There are two crucial assumptions in this hypothesis: (1) that self-respect has first priority among the primary goods, and (2) that self-respect has an intimate connection with basic liberties. The first assumption does not seem to present much problem. To establish it, it seems sufficient to examine how each primary good is characterized. If by its content, the primary good of self-respect is indeed more relevant with respect to rational life plans, then

is the logical deducibility of the principle itself from the conditions of the original position.

49 Rawls, op. cit. p.544.

Note here the fusion between self-respect and self-esteem. Rawls should probably be talking of 'self-respect' in the second sentence and not self-esteem.

it is indeed to be ranked prior to other primary goods. But there is one important consideration why such a move appears difficult or perhaps impossible. There are many quite different rational life plans, and unless one is able to identify something which is commonly crucial to every life plan and which requires self-respect, it seems rather vacuous for one to claim that self-respect is more relevant to life plans than every other primary good. Consider once more the life plan of a nuclear engineer and that of a clover leaf picker. What can these two plans have in common and how is self-respect a crucial requirement for such a common feature? It is not very clear how Rawls might respond to such a question.

The situation is even more complex with regard to the second assumption. A necessary or unique link must be shown to exist between the primary good of self-respect and the primary good of basic liberties. Here the second form of the argument for the equality of liberty principle becomes relevant. In the latter, Rawls assumes the supreme importance of self-respect as a logical premise for the conclusion that liberty is to take priority over other goods. In this form, an establishment of the priority of liberty serves causally as the most effective social basis available for self-respect. Henry Shue constructs an example of a similar logical structure as follows:

... the first priority is to assassinate the emperor; the best source of assassins available is discontented intellectuals; therefore, we must keep the intellectuals discontented. The supreme importance of achieving the end serves logically as support for the importance of taking a causally effective means to that end. (50)

In this example, once it has been established that the first priority is to assassinate the emperor, the next real task becomes that of supporting the claim that making the intellectuals discontented is the best means of accomplishing the assassination. Similarly, once we have assumed the correctness of Rawls' thesis that the most important primary good is self-respect, then the next and most important task is to show that equality of basic liberties is the most effective means of attaining the good of self-respect.

Rawls does not go into any detailed discussion as to how the primary good of self-respect is to be distributed. He may have presumed that being a primary good (the most important primary good at that) only an equal distribution of self-respect would be acceptable to the rational persons of the original position. Perhaps Rawls considers this point too obvious to require any extended argument. Yet, if this is indeed a correct reading of him, then there is a problem on hand. It might be the case that for people who already have a sense of justice, it would be rather trivial to argue for an equal distribution of self-respect. But the thin theory of the good is used to derive the principles of justice. Hence an appeal cannot be made to the sense of justice here. The parties are rational to the extent that they know that having more of the primary goods than less is

50 Henry Shue, "Liberty and Self-respect", Ethics, 85, 1974-5, p.147.

the most effective means of ensuring the realization of their life plans. They also know that whatever else they might want, they must want the primary good of self-respect. But how does this type of rationality lead to an equal distribution of self-respect? It would seem that one of the two initial assumptions in the argument for the priority of liberty is at least vulnerable. As stated earlier, the first assumption concerning the supreme importance of self-respect could be established if the content or conditions of self-respect can indeed be seen to be more relevant to all conceivable rational life plans than other primary goods. Perhaps Rawls might be able to do this in a non-trivial way. But the second assumption concerning the guarantee of an equal distribution of self-respect does not necessarily follow from the first assumption. The latter must be argued for independently. But Rawls does not appear to have done so.

However, assuming that these two assumptions are correct and that Rawls can indeed argue for them, there is still a third assumption in the inference from self-respect to the principle of equal liberty. If self-respect must be guaranteed an equal distribution, and if there exists an acceptable means for guaranteeing it, then this means must be undertaken. The equal distribution of the social basis of self-respect is implied in this hypothesis. If self-respect is supreme among the primary goods and if it must be equally distributed, then whatever serves as its social basis must also be equally distributed. This is the third major assumption in Rawls'

argument. Again, it seems so self-evident that Rawls must have thought it unnecessary to demonstrate its reasonableness.

...the basis for self-esteem is a just society ... the publicly affirmed distribution of fundamental rights and liberties. And this distribution being equal, everyone has a similar and secure status when they meet to conduct the common affairs of the wider society. (51)

Let us suppose that the inference in the third assumption is correct: since self-respect is to be equal whatever serves as the social basis for self-respect must also be equal. It does not obviously follow from this that liberty is indeed the social basis of self-respect. Why is liberty the social basis of self-respect and not any other primary good such as wealth?

Rawls indeed considers equal wealth as a possible alternative to equal liberty. But he rejects the former in favour of liberty. As he says, the basis for self-esteem is a just society; not "one's income share but the publicly affirmed distribution of fundamental rights and liberties." (52) He invokes two considerations in support of this conclusion. First he says :

Suppose... that how one is valued by others depends upon one's relative place in the distribution of income and wealth. In this case, having a higher status implies having more material means than a larger fraction of society. Thus, not everyone can have the highest status, and to improve one person's position is to lower that of someone else. The means of status, so to speak, are fixed, and each man's gain is impossible. Clearly this situation is a great misfortune. Persons are set at odds with one another in the pursuit of their self-esteem. (53)

Rawls' argument seems quite attractive at first sight especially if we recall some of the common features of the capitalist society. With its philosophy of "free enterprise" (at least as far as the official ideology goes) the individual can legitimately acquire as much wealth as he pleases even if such acquisitions is to the detriment of the society at large. The very notion of "competition" suggests winners as well as losers. Hence, it would be difficult, as Rawls says, to expect individuals to co-operate in increasing the conditions of self-respect, if the latter is based on personal wealth and income. Why would anyone want to co-operate with others to improve their highest status, if by so doing one's position or self-esteem is lowered? If self-esteem is based on such fixed elements as wealth or income, it follows, as Rawls points out that, a gain in self-respect for one person is a loss in self-respect for another.

Given the pre-eminence of this primary good, the parties in the original position surely do not want to find themselves so opposed. (54)

Assuming that Rawls is indeed correct, one might yet ask why it must come to such a cut-throat competition for more self-respect. Rawls' interest no doubt is to discover a society which provides equal self-respect: a society in which individual persons would view one another as equal members of a moral kingdom of ends. But here the difficulty raised earlier becomes

52 Ibid. p.545.

53 Ibid.

54 Ibid.

more pertinent. Why would men with the type of motivation Rawls gives them in the original position be interested in such a society? For instance, why would rational self-interested men who are neither envious of one another nor inclined towards one another, want to be equal with one another in the distribution of self-respect? Why should it matter to them individually whether the other person has self-respect or not? Rawls might respond to these questions by saying that, because these persons do not know what the natural lottery would give them in real life after the veil of ignorance must have been drawn, it would be rational for them to desire equality of self-respect. If self-respect is distributed equally everyone would at least have one of the necessary means for attaining his life plan. This, Rawls might say, is only a matter of prudence.

Perhaps an appeal to prudential consideration may indeed be helpful. However, for such an appeal to be successful, it must be shown why these persons, at least some of them, would not want to take some risk. Rawls must show why they must have total aversion to risk. Even though they know that self-respect is a necessary condition for the realization of their life plans, they might still choose a model of distribution that allows everyone to compete for whatever amount of self-respect he would get. Since they are totally uncaring about one another, it should not bother anyone in particular even if they all know that some would have more of self-respect while others would not have enough. Rawls rules out the possibility of risk

taking in the original position. But it will be argued in Chapter three that there is nothing necessarily irrational about some risk taking. And if this argument succeeds, it is not plainly evident that the men in the original position will (because they are rational), accept an equal distribution of the supreme good of self-respect just because they see the latter as supremely important to everyone's life plan.

However, assuming that these men could be made to see the need of accepting an unequal distribution of self-respect by appealing to some considerations, say the Hobbesian type, it is not certain why an equal distribution of self-respect cannot be attained from a purely economic stand-point. For example, let us suppose that people's relative degree of self-respect is based on their relative economic position -- this in fact seem to be a valid supposition owing to the manner Rawls characterizes the notion of self-respect, equating it with self-esteem -- it is yet possible to attain equal self-respect provided there is equal wealth. In which case the members of a well-ordered society would co-operate with one another to make sure that wealth was kept equally distributed, and if conditions permit, they would see to it that the general level of wealth was increased for all.

The idea of equal wealth does not appeal to Rawls. His idea of a well-ordered society is not a society without inequalities; but, rather, a society in which all inequalities are justified to the extent that they favour the least advantaged.

Hence the idea of striving for equal distribution of wealth in order to ensure an equal distribution of self-respect is not acceptable to Rawls.

Moreover, as I mentioned in the discussion of envy, if the means of providing a good are indeed fixed, and cannot be enlarged by co-operation, then justice seems to require equal shares, other things the same. But an equal division of all primary goods is irrational in view of the possibility of bettering everyone's circumstances by accepting certain inequalities. Thus the best solution is to support the primary good of self-respect as far as possible by the assignment of the basic liberties that can indeed be made equal, defining the same status for all. At the same time, distributive justice as frequently understood, justice in the relative shares of material means, is relegated to a subordinate place. (56)

This passage introduces a qualitative distinction with regard to the primary goods. While all the primary goods are indispensable means for the realization of rational life plans, some primary goods are by nature "fixed and cannot be enlarged by co-operation". Others are capable of being enlarged by co-operation. The constraint of equal distribution only applies to the former and not the latter. In order to better everyone's circumstances an inequality in the distribution of the latter must be acceptable. Thus wealth cannot be a social basis for the primary good of self-respect. While it may be rational for a society to pursue increased wealth up to some point, it is not necessary for it to seek equal distribution of wealth. An inequality in the distribution of wealth can be to everyone's advantage including the least advantage.

In fact as the conditions of civilization improve in the society the value of economic advantages diminishes in favour of liberty.

Now the basis of the priority of liberty is roughly as follows: as the conditions of civilization improve, the marginal significance for our good of further economic and social advantages diminishes relative to the interests of liberty, which become stronger as conditions for the exercise of the equal freedoms are more fully realized. Beyond some point it becomes and then remains irrational from the standpoint of the original position to acknowledge a lesser liberty for the sake of greater material means and amenities of office. (57)

The inference here is rather straight-forward. Since the value of wealth can diminish and that of liberty cannot, then equal liberty ought to rank prior as the basis of the primary good of self-respect.

There are two important points that we need to reflect upon in this argument. The first is the intimate connection between power and wealth. Judging from the way things are in our society, it is fair to say that power and wealth go almost hand in hand. Very often the amount of power an individual has is in direct proportion to the amount of wealth he controls. But from Rawls' very definition, self-respect is also directly linked to power. Does this factor not suggest that the condition making self-respect stable cannot be achieved in our society? Obviously, Rawls wants to break the psychological connection between wealth and self-respect: A person's self-respect must not be based on his relative de-

gree of wealth. But what chances does he have to succeed, if indeed there is such a strong link between power and wealth? To what extent can we talk of equal self-respect if there is inequality of power?

The second point that we need to consider is the deep psychological assumption that Rawls is making in this argument. Rawls assumes that material inequalities do have positive value in a just or well-ordered society. They serve as an incentive to some citizens to optimize their talents and abilities in ways which will improve the position of all citizens, especially the least advantaged. Hence the basic structure of a well-ordered society must be designed in such a way as to encourage or at least permit certain inequalities.

If there are inequalities in the basic structure that work to make everyone better off in comparison with the benchmark of initial equality, why not permit them? The immediate gain which a greater equality might allow can be regarded as intelligently invested in view of its future return. (58)

The claim that material inequality works to make everyone better off could be challenged on empirical grounds. Perhaps if we limit our consideration to the more industrialized societies of the Western world, this claim might seem less difficult to justify. In these societies people are generally encouraged to work harder in view of the material benefits they hope to receive. In the final analysis some benefits manage to trickle

58 Ibid. p.151.

down to the least advantaged through the hard work of the better talented. However, when we come to examine the situation in the less developed societies, especially the so called third world countries, it is not so evident that material inequalities always work to make everyone better off. The regular pattern in most of these societies is that the richer some people are, the poorer others become. In many African societies the term "middle class" hardly conveys anything. The polarity is always between the very rich and the very poor. Thus it cannot be said quite categorically that material inequality by itself always work to everyone's advantage.

Apart from the issue of validating the assertion that material inequality always work for everyone's advantage, there is still another important question concerning Rawls' notion of liberty. As seen earlier, Rawls is prepared to allow some inequality in wealth but not in liberty. But is wealth not a primary good? Why can some inequality of liberty not also be allowed in order to create incentives that will in turn benefit everyone? For instance, it is an accepted practice in several Western societies that citizens of certain rank are exempted from conscription into national military service. Why can such a practice not be justified along the same line as that of inequality of wealth? One reason why Rawls would not allow inequality of liberty lies in his very conception of liberty. Liberty, for him, is to be classed among those goods or means

or providing goods that are "indeed fixed and cannot be enlarged by co-operation". Thus, while a society may legitimately aspire to increase the level of its wealth by a cooperative venture, it cannot do likewise with regard to liberty. Liberty cannot be jointly pursued with a view to increasing or enlarging it. In Rawls' belief, an increase in liberty for one person is necessarily a net loss in liberty for another person. Liberty is always liberty from others. Hence it would be irrational for anyone to offer more liberty to someone else as an incentive, since this would mean a reduction of liberty for himself, and hence a defeat in advance of the very purpose of offering the incentive. The system of liberties is a zero-sum game.

However, this notion of liberty is by no means self-evident. In concrete everyday interactions, people make trade-offs in liberties in order to attain some other goods. It is not so apparent that whenever such trade-offs are made, the agents involved are necessarily irrational. Rawls' notion of liberty will require a more careful analysis before it can be pitched against the concept of wealth as a competing candidate for the social basis of the primary good of self-respect.

From these remarks, it seems that the inference from the primary good of self-respect to the principle of equality of basic liberties is not as successful as the two earlier inferences. The latter presents two main difficulties: The very assumption that the parties in the original position will desire

an equality of self-respect has not been adequately argued for, nor has a necessary connection been established between the primary good of self-respect and equality of basic liberties. If it were not for the fact that Rawls has a peculiar conception of liberty, it would be possible to reject equal basic liberties in terms of considerations similar to those in virtue of which he rejects equality of wealth. And with power being so intimately connected with wealth, inequality of wealth can only lead to inequality of power. And, inequality of power does not seem very compatible with the notion of equal self-respect.

Perhaps these objections could be met in a very important way, or even be shown to be invalid. Whichever way, the importance of the logical derivability of the two principles of justice as fairness need not be exaggerated. It is doubtful whether Rawls himself attaches too much importance to this exercise. Rather, the more crucial issue for him, as it is for several of his critics, is whether the principles emerging from the original position can adequately articulate our sense of justice. In other words, why must we regard the principles that might be chosen under the conditions of the original position as principles of justice? We conclude the inquiry in this chapter with this question. The discussion will focus on the restrictions on knowledge at the original position.

THE CHOICE OF MORAL PRINCIPLES FROM BEHIND A VEIL OF IGNORANCE.

One of Rawls' main tasks in A Theory of Justice, as stated earlier, is to evolve a conception of justice that can effectively regulate a well-ordered society. A conception of justice that truly reflects our nature as free and equal moral persons, and which rational persons would mutually agree to both as a basis of social co-operation and limiting constraints on individuals in the pursuit of their ends and purposes.

The main source of challenge for Rawls here is his conception of a well-ordered society. Rawls' notion of a well-ordered society articulates an ideal of social life that is strongly pluralistic. In his well-ordered society, individual persons and groups have, and view themselves as having fundamental aims and interests in the name of which they make legitimate claims on one another. Their conceptions of the good incline them in contrary directions. And, furthermore, they have opposing basic beliefs (religious, philosophical, and moral) and their ways of assessing evidence and arguments are different in many essential cases. Despite the divergence of fundamental interests and ends of its members, and their variety of opposing and incompatible basic beliefs, the society is still well-ordered because it is effectively regulated by a public conception of justice. That is:

1. Everyone accepts and knows that others accept the same principles (the same conception) of justice.

2. Basic social institutions and their arrangement into one scheme (the basic structure of society) satisfy and are with reason believed by everyone to satisfy these principles.
3. The public conception of justice is founded on reasonable beliefs that have been established by generally accepted methods of inquiry. (59)

What Rawls must show here is how individual persons who have their own ~~disparate~~ ends and purposes, and who have no inclination towards one another can yet agree mutually on certain basic principles of justice to govern their institutions and themselves. To do this Rawls introduces a veil of ignorance at his original position. The general idea is to control all the natural and social contingencies that may tend to bias individuals in the choice of moral principles. Rawls wants to strip the contractors of all morally irrelevant contingencies that would lead some of them to press for special advantages, or give some of them special bargaining power. His strategy is to induce the contractors to reason from the same basic premises. By freeing them from their individual personal biases and by making them to reason from an identical standpoint, Rawls expects that they would inevitably arrive at a unanimous choice of moral principles.

Thus, the veil of ignorance is an attempt by Rawls to secure a unanimous choice of the principles of fairness by removing the effects of specific contingencies which put men at

59 John Rawls. "REPLY TO ALEXANDER & MUSGRAVE".
Journal of Economics, Vol.88, 1974. pp. 634-5.

odds and tempt them to exploit social and natural circumstances to their own advantage. The parties do not know their own social situation, their natural endowments or even their individual life plans together with their conceptions of the good. And, in order to secure fairness between generations, they are also deprived of the knowledge of the particular political, economic and cultural circumstances of their society. Under such a thick veil of ignorance, Rawls expects that they would opt for principles of toleration. According to him, this is the only rational thing for them to do, if they are yet to safeguard their individual ideals and values. They would insist that all the basic opportunities which are necessary for the realization of the various ideals of life be equally distributed.

Before considering the effects of its specific restrictions, we need to inquire into the very rationale of the veil of ignorance itself. What is the justificatory force of this constraint? Precisely how can an agreement reached under conditions of ignorance, as a representation of consent, by itself provide adequate moral justification for the principles of justice? Thomas Nagel makes some important observations on this question. He says:

When one justifies a policy on the ground that the affected parties would have (or even have) agreed to it, much depends on the reasons for their agreement. If it is motivated by ignorance or fear or helplessness or a defective sense of what is reasonable, then actual or possible prior agreement does not sanction anything. In other cases, prior agreement for the right reasons can be obtained or presumed, but it is

not the agreement that justifies what has been agreed to, but rather whatever justifies the agreement itself. If, for example, certain principles would be agreed to because they are just, that cannot be what makes them just. In many cases the appeal to hypothetical prior agreement is actually of this character. It is not a final justification, not a mark of respect for autonomy, but merely a way of recalling someone to the kind of moral judgments he would make in the absence of distorting influences derived from his special situation. (60)

Nagel's observations are obviously correct, but it is doubtful whether Rawls himself would want to challenge them. Rawls' appeal to the hypothetical prior agreement at the original position cannot be construed as a final justification of the principles of justice as fairness. Rawls knows, as we would expect, that the mere fact that certain principles would be chosen because they are just, is not the reason that makes them just. Actual or hypothetical agreement becomes a justificatory factor only when the parties concerned have made a prior commitment to treat one another as certain reasons would incline them to. Thus the veil of ignorance is no more than a device to create the circumstances of consent that would generate the kind of moral judgment that rational persons would ordinarily agree to make, in the absence of distorting influences which derive from their particular and special life situations. But then, does Rawls' account not tend to lose some of its distinctive features in view of such a rendering? This need not be the case, especially if we recall an observation

60 Thomas Nagel, "Rawls on Justice", in Reading Rawls, op. cit., p.5.

made in the previous chapter concerning two distinctive types of arguments at the original position: the argument for justice as distinct from the argument for fairness. Whereas the former tries to discover what principles might yield just distributions under the conditions of the original position, the latter tries to articulate the conditions under which the kind of moral judgments of fairness that rational persons would ordinarily make can emerge. Rawls' idea of "pure procedural justice" is very crucial in this respect. As opposed to the Cartesian approach, Rawls is not trying to discover any new and eternal principles of justice. His purpose is to construct a fair procedure from which the principles of justice would emerge. The idea of the original position, he says, "is to set up a fair procedure so that any principles agreed to will be just".⁽⁶¹⁾ His central assumption may be put as follows: if there is any procedure for distribution that rational persons would see and accept unanimously as fair, only such a procedure may yield the principles of justice.

It follows therefore, that even though the particular interests of the rational contractors may not coincide; and, even if they may not agree on how the particular advantages which may subsequently derive from their co-operation are to be distributed, they may yet agree unanimously on one basic

61 Rawls, A Theory of Justice, p.136.

issue: that is, a fair procedure for adjudicating their conflicts. They may agree that the ground rules of their procedure for distribution treat each party equally in relevant respects.⁽⁶²⁾ To ensure such a unanimity, a veil of ignorance must be drawn over the particular interests as well as the individual capacities of the contractors. By trying to limit himself to reasons that all men have in common, especially as moral persons, and by omitting those considerations which would lead them to select different principles and incompatible institutional structures, Rawls may be able to make a legitimate appeal to the notion of prior consent. In fact, since he is primarily interested in deriving moral conclusions (principles), the contractors will not be in any disadvantage if they do not have access to any non-moral facts. Thus it would seem that if the idea of the veil of ignorance is put in proper perspective it need not raise any serious difficulty.

However, while the idea of the veil of ignorance may thus be defensible as a way of ensuring that the parties are on an equal footing (they do not possess any information which lead them to seek undue advantage on morally irrelevant grounds),

62 There is the question, of course, whether any procedure can in fact treat everyone equally in respect of everyone's relevant interests? What would be the criteria of assessing the relevance of everyone's interests? It is not immediately evident how issues such as these could be satisfactorily resolved.

the question yet remains whether all the information it rules out are indeed morally irrelevant. Two restrictions at the original position are particularly relevant to this question: the denial of particular conceptions of the good and the denial of moral conceptions and sentiments. We examine both of these beginning with the first.

While such items as race, sex, parentage or natural endowments may indeed be irrelevant as grounds for moral judgments, it is not at all obvious why the particular conceptions of the good of the contractors should also be considered irrelevant. Should any of the contractors favour certain principles which could advance his particular conception of the good, how does this necessarily imply that he is seeking special advantages for himself over all others? For one thing, none of the contractors know who he is in the society. Hence in choosing his set of principles, he is at the same time seeking to advance the good as conceived by him for everyone in the society including himself. If the conception of the good is not equated with a mere system of tastes or preferences, then there is nothing so odd or illegitimate about someone wishing to advance the good as conceived by him for everyone in the society.

The inclusion of this restriction at the original position raises a very important question concerning Rawls' theory of the good. Admittedly, people's different conceptions of the good often lead to disagreement and conflict in the society. But does this reflect the fact that the good is of its very

nature controversial and people cannot come to a mutual agreement about it, or that people generally do not have adequate conception of the good? Or even if it is the case that people's individual conceptions of the good are always distorted because of their individual psychologies, this factor should not be a problem since Rawls has purged his parties of their individual psychologies. Why would such "rational" persons yet differ on a theory of the good such as Rawls'. An account based on what Rawls himself takes to be objective principles of rational choice? The answer to this question is rather obvious. Rawls cannot appeal to the full theory of the good at the original position since the latter already contains some considerations about the principles of justice he is trying to derive. The contractors cannot be presumed to have a conception of the good which already includes the principles of justice which the original position is supposed to yield. But could this then not be the real reason for the exclusion of the individual conceptions of the good. Namely, that the conceptions might include considerations of morality. But this is not to say that conceptions of the good are morally irrelevant or that rational persons cannot come to an agreement about them. Rawls' exclusion of individual conceptions of the good is motivated primarily by the consideration that these cannot function suitably as premises in view of the approach he wants to follow. While this move by itself may be permissible from a logical standpoint -- a logician is normally at liberty to introduce whatever pre-

mises he considers suitable for the particular theorem he wants to prove, provided he discharges these premises in the end -- it is doubtful whether it would work for Rawls. The features of the original position, as he says, are peculiar to morality. To make good on this claim, his assumptions must be consistent with morality in general and not just his conception of morality. This point will be clarified in the final chapter of this study.

Meanwhile, assuming that Rawls is indeed correct in ruling out individual conceptions of the good or that the questions raised in the above paragraph can be effectively countered, there is still another restriction under the veil of ignorance that has caused considerable resistance. The latter is the denial of substantive moral conceptions or moral sentiments to the contractors. How acceptable is this restriction?

The rationale for this restriction stems largely from what Rawls considers to be the circumstances in which questions of justice arise. These are circumstances of moderate scarcity and conflict of interest. Should there be an abundance of human goods, perfect altruism and social harmony, there would hardly be any disputes among men. Hence it might not be necessary to have principles of justice as there would be no conflict of interests to adjudicate. Thus, in respect to the original position, Rawls says that "when it is supposed that the parties are severely disinterested, and are not willing to have their interests sacrificed to the others, the intention is to

express men's conduct and motives in cases where questions of justice arise"⁽⁶³⁾.

Rawls' observation seems obviously true. A society in which there is a strong diversity of interests, and in which conditions of moderate scarcity exist, would require principles of justice to regulate the conflicting claims of its members. But this fact does not imply that justice would not be relevant in a society in which people have substantive moral notions. For instance, a society of angels, where there is perfect altruism and social harmony; the question of just regulation may yet arise especially if conditions of moderate scarcity prevail. As Brian Barry recalls, people in real life disagree on moral grounds "because they have conflicting interests; (but) if we put them into a situation where these conflicting interests cannot influence them, they can reach agreement"⁽⁶⁴⁾. One gets the impression at first that the strategy of the original position is designed precisely to attain such a situation: A situation where the conflicting interests of the contractors cannot influence them because these interests are unknown to them. But then why should Rawls also extend the veil of ignorance to cover substantive moral notions? If Barry's observation is correct, and this appears to be so, can the contractors not still agree on a choice of moral principles even though they are aware of their moral notions, provided their conflicting interests are unknown to them?

63 Rawls, op. cit., p.129.

It seems rather obvious that without ruling out moral notions at the original position, Rawls' moral geometry is in jeopardy. If the parties were allowed to retain their moral notions, it becomes difficult for Rawls to generate a deductive formula. For instance, if he were to say that in the absence of self-interest the original contractors would agree on the principles of justice as fairness, such a statement would simply be assertive. It does not even have the appearance of a deduction.

However, in its deductive form, Rawls' formula has not successfully shown why principles adopted from the point of view of self-interest ought to be accepted from the point of view of morality. Even if the contractors are said to be rational, since they are not altruistic, they do not have the capacity to act from the moral point of view. As stated earlier, to be able to act from the moral point of view, an agent must be able to see the good of the other person as a sufficient reason and motive for action. Indeed, Rawls' men may still choose the two principles of justice as fairness as means of regulating their competing claims. They may desire equality of "self-respect" as a means of attaining their individual ends. But while such a choice or desire may be considered rational (here we would be talking of the rationality of means-to-end), the contractors cannot be described as moral persons nor can their choice be considered a moral choice. As

we will come to see in later chapters, an amoralist can under certain circumstances, rationally appeal to the principles of justice as fairness in the interest of his own private ends or the ends of his friends or members of his own class. But this would not make him a moral person.

It would seem from these remarks that the original position runs into serious difficulties as a theory of moral justification. As seen earlier, Rawls' contractarian method seeks to provide a device for displaying and deciding what moral principles or system of ethics would be adopted by rational men in a position where they would have to be mutually disinterested. While this is indeed a very interesting task, it is almost an impossible one. These rational men are not just capable of making a moral decision because the very position they occupy, (the position of prudent maximizers) cannot provide a basis for moral decisions.

SUMMARY AND CONCLUSION

The concept of justice and fairness might seem identical at first sight. Hence we might not see the need either to distinguish between them or assert one as being more fundamental than the other. According to Rawls, this approach is quite mistaken. In his theory entitled Justice as Fairness, he tries to show the concept of fairness as the fundamental idea in the concept of justice. From this point of view, Rawls sets out

to provide an analysis of the concept of justice based on the ideas of the social contract. He argues that utilitarianism is unable to account for this aspect of justice.

In the beginning of his analysis, Rawls states two principles which specify the conception of justice as fairness. He then considers the circumstances and conditions under which these principles can be said to arise. This chapter dealt mainly with these conditions which Rawls designates as the original position. The two principles defining justice as fairness, as Rawls asserts, are quite familiar. However, by invoking the notion of the original position, Rawls attempts to show the basic structure of these principles and how they conform to our ordinary intuitions about just distribution. Hence the concept of the original position could be seen to play two major roles in Justice as Fairness (an analytic role and a justificatory role). As an analytic device, it purports to help us understand the degree to which our conception of justice depends on various formal requirements such as the conditions of generality, finality, and publicity that apply to the principles. It also provides a concrete model for reducing the complex problem involved in the social choice of principles of justice, to a more manageable problem by way of the rational individual choice of principles.

In its justificatory role, Rawls seeks to construct the original position such that the choice of the principles

emerging from it would constitute a justification for these principles. He restricts the grounds of choice by introducing certain constraints on knowledge and on motivation. The contractors are stripped of all morally irrelevant contingencies that could lead some parties to press for special advantages, or give some of them special bargaining power. And, in order to articulate what Rawls considers to be the circumstances in which questions of justice can only arise (circumstances in which conflict of interests prevails), the contractors are also denied of all substantive moral notions and sentiments including their individual conceptions of the good. Thus the original position presents a picture of mutually self-interested persons who must choose principles of just distribution from behind a veil of ignorance. Since they reason from the same basic premises, Rawls expects that they would inevitably arrive at the two principles of justice as fairness. How successful is the use of the original position in these two capacities?

It should seem evident from the lengthy discussion of the original position that this question cannot be answered by a straight "yes" or "no". In fairness to Rawls, the notion of the original position does indeed capture some important conditions (necessary conditions) that must be entertained in evolving principles of justice or morality. For instance, such conditions as generality, finality, and publicity are

essential ingredients toward the construction of any principle in normative ethics. However, Rawls' interpretation of the original position runs into serious problems which are not easily resolvable. In its analytic role, there is the difficulty, logical or otherwise, of isolating the most reasonable interpretation of the initial situation that yields a particular conception of justice. Even if it is possible to construct an interpretation whereby the principles of justice as fairness can be seen to follow from the original position, it is logically possible to derive the same principles from a different set of premises which could be shown to be as reasonable as the ones used by Rawls. And granting this possibility, the original position cannot be considered unique to the principles of justice as fairness. To overcome this difficulty, the premise set in which Rawls is interested must be logically equivalent to his principles. But this is not so in his account. Hence it could be entertained that the principles of justice as fairness can be derived from a set of premises other than the ones assumed by Rawls. By the same token, there is, at least, a theoretical possibility of deriving the principle of average utility from a set of premises different from those employed by Rawls. Hence Rawls' rejection of the latter principle in the light of his own interpretation of its initial situation is rather inconclusive.

The notion of "reasonableness" also presents some problem in the analytical use of the original position. As

noted earlier, the notion of "reasonableness" is not one in which a consensus can easily be reached even among rational persons. The criteria of reasonableness specified by Rawls are by no means conclusive. In choosing among conflicting conceptions of justice rational persons may never agree on the most reasonable interpretation of the initial situation.

In its justificatory role the notion of the original position runs into serious problems in view of the restrictions imposed by Rawls. While prior consent may indeed be cited as legitimate grounds for abiding by the terms of a contract, it is not at all obvious whether an agreement based on ignorance can have any justificatory force. Why should a hypothetical choice under conditions of ignorance, as a representation of consent, by itself provide a moral justification for outcomes that could not be unanimously agreed to if they were known in advance? However, granting that Rawls' attempt to eliminate morally irrelevant contingencies at the original position could be made plausible, why should such vital questions as the particular conceptions of the good be regarded as morally irrelevant? This question calls for a close examination of Rawls' conception of the good. This task will be undertaken in part two of this study.

Apart from helping to create the appropriate circumstances in which questions of justice arise, the denial of substantive moral notions to the men of the original position is

very crucial to Rawls' deductive formula. For instance, if Rawls were to say that in the absence of self-interest, masses of people would agree to the principles of justice as fairness, such a statement would be a mere categorical assertion and not a deduction. But, in its deductive form, the statement fails to show why principles adopted by self-interested (non-altruistic) persons in the original position ought to be accepted as principles of justice.

Perhaps with some extended arguments, some or all of these difficulties could be overcome. However, until we are confronted with such arguments and their soundness is established or at least shown to be plausible, the concept of the original position remains problematic and justice as fairness and its principles appear at least to be essentially contestable. This survey will now continue with an examination of Rawls' moral methodology.

CHAPTER TWO

RAWLS' MORAL EPISTEMOLOGY

To present an adequate picture of the essential framework of A Theory of Justice, one needs to go beyond the original position and the two principles of justice which Rawls tries to derive from it. With respect to the latter, Rawls states rather explicitly in his concluding remarks, that the two principles are "contingent in the sense that they are chosen in the original position in the light of general facts"⁽¹⁾. If we must look for "likely candidates for necessary moral truths", he suggests that we proceed in the direction of "the conditions imposed on the adoption of principles"⁽²⁾. Thus for Rawls, what is uniquely important in A Theory of Justice is not so much the plausibility of the two principles of justice he articulates in it, as the method by which he arrives at these principles. This holds true also for the basic principles regulating other aspects of morality. The question of methodology ranks prior to the choice of substantive principles. This is so because our very choice of principles may well be determined by the criterion of choice we appeal to.

1 Rawls, A Theory of Justice, p.578.

2 Ibid.

From the results of the analysis in the preceding chapter, it does not seem to be the case that Rawls' characterization of the original position is such that any principles which it would be rational to choose in this situation are bound to be principles of justice or acceptable from a moral point of view. The assumptions in the original position are not peculiar to moral theory. And, for the most part, they presuppose, for their plausibility, the acceptance of the conception of justice Rawls wants to evolve. However, Rawls still believes that he has an effective answer to the challenges that might be raised against his strategy of justification in the original position. The two principles of justice he formulates, his specification of the original position, and in fact the entire conception of justice as fairness, are to be considered adequately justified if they could be shown to accord with our "considered judgments" of justice. In other words, justice as fairness would pass for an adequate moral theory, if the substantive principles deriving from it, when conjoined to our beliefs and knowledge of circumstances, will match our common intuitive moral convictions in particular instances.

Rawls refers to this method of testing the validity of an ethical theory as the method of "reflective equilibrium". This method offers a criterion for deciding which of particular rival moral theories is to be preferred. The purpose of it is

to see how the principles we may be considering match with our most firmly fixed and well considered moral beliefs. This appeal to our fixed and firm moral beliefs (moral intuitions as he calls them) is, according to Rawls, unavoidable in moral reasoning. It is a method which goes back to Aristotle, and can be followed down through the classical writers, at least to Sidgwick. Rawls thinks that, if we appeal to this method, we will come to accept justice as fairness as "a theory of our moral judgments in reflective equilibrium"⁽³⁾.

Perhaps it might not be necessary to go back to Aristotle or any of the classical writers to try to discover how the method of reflective equilibrium is used, if it has indeed been used at all. Rather what will be done here is to examine the adequacy of the method itself as a test for the validity of moral theories. Several critics of Rawls' theory, notably Hare, Peter Singer and Steven Lukes, object vehemently to this method. They maintain that Rawls' appeal to it exposes his theory to a kind of "radical subjectivism". They contend that the "facts" with which Rawls' theory has to square are nothing but people's opinions. Hence the statements of Rawls' principles cannot be objective or sound because they rely ultimately on "received opinions". Other

3 Rawls, Ibid. p.120.

critics challenge the method on the ground that our ordinary moral judgments (common-sense moral convictions) are much less firmly based than scientific observations against which scientific theories are tested. They maintain that, in contrast with science, there is no highly developed procedure that would enable the theorist to go beyond common sense, and to reject what merely appears to be morally right or wrong.

An attempt will be made to investigate these criticisms in this chapter. It would appear that the method of reflective equilibrium does offer a useful procedure for testing the adequacy of a moral theory. As Aharon F. Kleinberger rightly observes, people's

...opinions, convictions and intuitive judgments are the kind of objective facts with which the social and behavioral sciences have to deal, and against which their theories have to be tested. And although it may be granted that people's opinions are less constant than, say, the motions of planets, and that the methods of observing the latter, this does by no means imply that the observational methods of the social and behavioral sciences are incapable of discriminating between real and merely apparent opinions, or that they cannot go beyond common sense (even though common sense or subjective attitudes may be the object of their research). (4)

Thus an outright rejection of people's moral convictions as one of the factors to be considered in testing the validity of a moral theory is inappropriate. Rather what we need to

4 Aharon F. Kleinberger, "The Social-Contract Strategy for Justification of Moral Principles", Journal of Moral Education, Vol.5, No.2, (1976), p.122.

do is to examine carefully the kind of theory that a theory of justice or morality is supposed to be, and then proceed to determine whether or not common convictions and judgments of justice or morality are an appropriate test for that particular kind of theory. Accordingly, we shall first examine Rawls' notion of a theory of justice in the context of his conception of the nature and scope of moral philosophy.

MORAL PHILOSOPHY: AN ATTEMPT TO EXPLAIN OUR MORAL CAPACITY

Following the general point of view in his "Outline of a Decision Procedure for Ethics"⁽⁵⁾, Rawls broadly characterizes moral philosophy "as the attempt to describe our moral capacity", and his theory of justice as an attempt to describe our sense of justice"⁽⁶⁾. The term "moral capacity" is a key element in this characterization. By it, Rawls means the extraordinarily complex skill which we normally develop at a certain age. This capacity enables us both to judge actions as right and wrong; and to act in accordance with our judgments while expecting others to do likewise.⁽⁷⁾ Rawls admits that the task of moral philosophy is both complex and difficult. As a systematic and theory-laden

5 Rawls, John, "Outline of a Decision Procedure for Ethics", The Philosophical Review, 60, (1951), pp.177-97.

6 Rawls, op. cit., p.46.

7 Ibid.

description of our moral capacity, it requires, above all, "a formulation of a set of principles which, when conjoined to our beliefs and the knowledge of circumstances, would lead us to make these judgments with their supporting reasons, were we to apply these principles conscientiously and intelligently"⁽⁸⁾. Thus, in the process of a systematic probe into the moral capacity, we encounter those principles which regulate our moral judgments and convictions. Hence the proper function of the moral philosopher consists mainly in trying to isolate such principles that could form part of the premises from which we can derive specific moral conclusions. Such conclusions should necessarily conform with the great majority of our considered judgments. In A Theory of Justice therefore, Rawls may be seen as trying to characterize "our moral sensibility when the everyday judgments we do make are in accordance with its principles"⁽⁹⁾. His basic assumption is that "we do not understand our sense of justice until we know in some systematic way covering a wide range of cases what these principles are"⁽¹⁰⁾.

Rawls claims that his conception of moral philosophy is most effectively captured in Chomsky's analysis of the

8 Ibid.

9 Ibid.

10 Ibid.

problems of characterizing a native speaker's grammatical capacity. The linguist seeks "to characterize the ability to recognize well-formed sentences which make the same discriminations as the native speaker"⁽¹¹⁾. In a similar move, the moral philosopher attempts to depict our moral capacity. Rawls contends that despite the difference in their subject matter, both the linguist and the moral theorist go through an identical exercise: both of them must develop such theoretical constructs that would go beyond traditional grammar books and familiar common sense moral precepts respectively. He does not see how else they can attain any significant overview in their respective disciplines should they proceed otherwise. How valid is this analogy?

In his critical review of Rawls' theory, Thomas Nagel rejects Rawls' analogy on the grounds that the intuitions of native speakers are decisive as regards grammar in a way that our first order moral intuitions are not.

Whatever native speakers (of English) agree on is English, but whatever ordinary men agree in condemning is not necessarily wrong. Therefore, the intrinsic plausibility of an ethical theory can impel a change in our moral intuitions. Nothing corresponds to this in linguistics, where the final test of a theory is its ability to explain the data. (12)

11 Ibid, p.47.

12 Thomas Nagel, "Rawls On Justice", The Philosophical Review, 82, 1973, p.201.

Nagel's observation is obviously right. The native speaker's usage easily appeals to us as the ultimate decisive factor in a way in which the first order moral judgments of the ordinary man does not. Besides, as Bernard Williams points out, "there is no adequate analogy in the linguistic case for the situation of conflict and disharmony in our moral sentiments, where, necessarily, it matters to us in practice how the conflict is resolved"⁽¹³⁾ When an appeal is made to the native speaker's usage in the linguistic case the goal and purpose of such an appeal is to explain a linguistic fact. But a moral theory, at least a normative moral theory, does not merely seek to explain our moral sentiments. Very often the situation of conflict and disharmony in our moral sentiments is one of the main reasons that leads to a moral theory. Bernard Williams further points out that Rawls is rather unclear as to how such conflicts and disharmony relate to moral theory. His analogy to linguistics does not help us to understand it.

It would appear at first that the disanalogy by itself does not pose any serious threat to Rawls' conception of ethical theory. One might even suggest that Rawls does not intend his illustration to be anymore than a useful comparison to show

13 Bernard Williams, "The Moral View of Politics", The Listener, 3 June, 1976.

how ethics (in his conception) attempts to explain a distinctive capacity by evolving principles and complicated theoretical constructs. However, it is doubtful whether Rawls would accept such a weak interpretation of his analogy. The analogy is much more crucial for him in view of his ethical cognitivist position. Rawls wants to maintain a close similarity between a moral and a scientific theory. For him, both kinds of theories are similar to the extent that they can make truth claims (claims that can be checked for their truth and falsity). It is in this sense that Rawls sees his theory of justice as being descriptive. But should the analogy be taken in such a strong sense, then Rawls must address himself to the criticisms raised in the preceding paragraph. It is not so clear how he can overcome these difficulties.

Apart from the theoretical issues raised by the analogy between linguistics and moral theory, Rawls' description of the task of moral philosophy as an attempt to characterize our moral capacity also raises a very important question: In what way does such a characterization capture the normative aspect of moral theory? For instance, how is a normative moral theory similar to a scientific theory? Specifically, how does such a conception fit Rawls' own remark that "we should view a theory of justice as a guiding framework designed to focus our moral sensibilities..."? (14) The terms "focus" and "guiding

14 Ibid., pp. 51-52.

framework" in this text, indicate a leaning towards a normative standpoint. Also, in a subsequent rejection of the view that meta-ethics is the central concern of a moral theory, Rawls declares that "a moral theory characterizes a point of view from which policies are to be assessed"⁽¹⁵⁾. Now, how do we reconcile these contrasting viewpoints? Perhaps the issue involved here will become clearer if we recall the influence of Quine's wholistic approach in philosophy on Rawls' thinking. As Kai Nielsen points out, it is important to keep in mind that Rawls has been deeply influenced on foundational matters by his colleague Quine. Like Quine, he does not make sharp distinction between the analytical task of the explication of concepts and an examination of substantive matters.

...Rawls makes substantive claims, builds his account on contingent matters of fact, appeals to scientific theories, and takes it as his task -- and the task of moral philosophy -- to give an explanation of our moral capacities, including our capacities to make and defend our considered judgments of rightness and wrongness and goodness and badness. Where he gives explanations, 'explanation', Rawls contends, has the same sense that it has in science, the methodological approach is the same, and Rawls' preferred explanations are open, he claims, to tests of a similar sort to what we have in the empirical science of linguistics. (16)

15 Ibid, p.286.

16 Kai Nielsen, "Philosophic Method," International Philosophical Quarterly, Vol. xvi, No.3, Sept., 1976, pp.358-9.

If Rawls is operating from such a background, it should not come as a surprise then that he should decide to define the problems and scope of moral philosophy rather differently from his predecessors. Nonetheless, while the approach itself need not be faulted simply because it departs from the traditional ones, it is still very doubtful whether it can actually lead Rawls beyond what has been commonly referred to as descriptive ethics. And while descriptive ethics does have its own positive value in ethical theory, its role is rather limited. For one thing it cannot lead to the Archimedian point that Rawls desires; nor can it guarantee the type of justification which traditional moral philosophers have actively sought. It cannot provide a justificatory apparatus, as Nielsen put it, to assess moral codes, social structures and conceptions of society. (17)

Again, the point here is not that Rawls' conception of the nature of a theory of justice is incorrect or inadequate, or that common convictions and judgments of justice are inappropriate as a test for that kind of theory. Despite the difficulties raised in the preceding paragraphs, his conception of what kind of theory a theory of justice is supposed to be, may yet be plausible. Rather, the main issue here is that Rawls cannot do what he hopes to do within

17 Ibid, p.365.

the framework of his conception of a theory of justice. From the remarks noted earlier, Rawls evidently sees a theory of justice as a descriptive-explanatory theory. Hence, like any of the theories in the social and behavioral sciences, a theory of justice should be tested against people's actual judgments in particular cases. But such a theory cannot be prescriptive-justificatory all at the same time. It cannot provide rationally derived principles against which people's actual judgments are to be tested for correctness. Thus as Kleinberger rightly points out, Rawls' attempt to conflate these two types of theory is necessarily doomed. To support this claim, it might be useful at this point to examine Rawls' formulation of the equilibrium argument.

OUR CONSIDERED JUDGMENTS IN REFLECTIVE EQUILIBRIUM

For Rawls, our considered judgments are those judgments which display our moral capacities without distortion. This category of judgments will exclude judgments motivated by such contingencies as fear, ignorance, self-interest, lack of confidence and the like. By contrast, our considered judgments are the judgments we make "under conditions favourable for deliberation and judgment in general". By shuttling

back and forth between these particular judgments, our conception of the original situation and our proposed moral principles, we attain the state of reflective equilibrium. Rawls demonstrates how this state is attained in justice as fairness as follows:

According to the provisional aim of moral philosophy, one might say that justice as fairness is the hypothesis that the principles which would be chosen in the original position are identical with those that match our considered judgments and so these principles describe our sense of justice. But this interpretation is clearly over simplified. In describing our sense of justice an allowance must be made for the likelihood that considered judgments are no doubt subject to certain irregularities and distortions despite the fact that they are rendered under favourable circumstances. When a person is presented with an intuitively appealing account of his sense of justice, (one, say, which embodies various natural and reasonable presumptions) he may well revise his judgments to conform to its principles even though the theory does not fit his existing judgments exactly. He is especially likely to do this if he can find an explanation for the deviations which undermine his confidence in his original judgments, and if the conception presented yields a judgment which he finds he can accept. From the standpoint of moral philosophy, the best account of a person's sense of justice is not the one which fits his judgments prior to his examining any conception of justice, but rather the one which matches his judgments in reflective equilibrium. As we have seen, this state is one reached after a person has weighed various proposed conceptions and he has either revised his judgments to accord with one of them or held fast to his initial convictions (and the corresponding conception). (19)

18 Ibid. p.48.

19 Ibid.

Before turning specifically to the notion of reflective equilibrium, it would be necessary first to indicate how this passage answers one important charge that has been made against Rawls: the charge of Platonism. As mentioned in Chapter One, P.H. Nowell-Smith, taking a cue from the conditions of the original position and especially Rawls' talk of "moral geometry in the strict sense of the term", alleges that Rawls' approach in A Theory of Justice is Platonic and hence he argues that the theory cannot succeed. Perhaps Rawls may have indeed given this impression. His very attempt to attain an Archimedian point to judge the whole of morality sounds rather Platonic. However, a careful reading of his statement of the reflective equilibrium in the passage cited above does not support the tendency of platonism. The reflective state is always subject to change. The principles arrived at are at best contingent. They may give way to other principles in course of time as new facts emerge. Thus Rawls' position may in fact coincide more with what Nowell-Smith (20) himself properly characterizes as the "Aristotelian" approach. Those considered judgments that we cling to as both secure and reliable emerge after a due process of reasoned examination of facts coupled with a thorough self-awareness and an understanding of our situation. As Rawls emphasizes, we can make such moral

20 P.H. Nowell-Smith, op. cit., pp. 315-7.

judgments only if we "have the ability, opportunity and the
desire to reach a correct decision"⁽²¹⁾.

Perhaps it should be pointed out here too, that, although Rawls recognizes the role of intuition in our moral judgments, he cannot be associated with pluralistic intuitionism. He does not claim that the best account of a person's sense of justice or the sense of right and wrong is the one which states his firmest considered judgments prior to his examination of any systematic philosophical conception of morality. To arrive at the state of reflective equilibrium, several factors come into play. The agent must weigh the facts available to him against a plausible conception of morality. Also, he must have a clear conception of the original situation together with an understanding of his particular situation. In going through this process he may even be compelled to revise his existing moral judgments when he has considered these against the constraints of a particular moral theory.

With these brief remarks about actual or possible mis-conceptions about Rawls' approach, we now turn specifically to the notion of reflective equilibrium. The central issue to be considered is whether this device as it emerges in the passage quoted above does indeed yield an objective

21 Rawls, op. cit., p.48.

standard for testing the adequacy of a moral theory.

From Rawls' statement above, and the subsequent remarks thereafter, the notion of reflective equilibrium may be represented as evidence or a condition of validity for a moral theory. In which case, when a moral theory matches a set of considered judgments in reflective equilibrium only then can it be said to be a valid moral theory. But the meaning of the term "valid" is yet opaque in this context. As Peter Singer points out, it is not easy to tell whether the fact of a moral theory matching a set of considered moral judgments in reflective equilibrium is to be taken as evidence of the validity of the moral theory, or rather, whether the reflective equilibrium is itself a definition of the validity of the moral theory. If it is the latter, it then follows that the achievement of a stable reflective equilibrium is what Rawls means by "valid"⁽²²⁾. Hence all that is required in the establishment of the validity of a moral theory is the conjunction⁽²³⁾ of the theory itself with a set of considered judgments. Because of the constraint of

22 Peter Singer, op. cit., p.493.

23 I make this conclusion on a rather tentative note. It would seem that the third and final argument of Rawls' theory, the argument of congruence does have some important bearing on the issue of validity. But in view of the complex nature of Rawls' work, it is not easy to locate the place and the effect of this argument in the structure of his theory. However, this question will be closely examined in Part Two of this thesis when we come to examine the congruence argument.

equilibrium, it might be necessary to revise the considered judgment until a perfect fit is attained between theory and judgments. Although the considered judgments are supposedly made under favourable conditions, an intuitively appealing account of one's sense of justice may yet call for their revision until they conform to the principles of such an account. Here the point of the analogy between linguistics and moral theory becomes very relevant. For the linguist to attain a correct theory, his construction must fit the native speaker's linguistic intuitions perfectly. It may be necessary for him to revise whatever conclusions he may have reached earlier in order to attain a perfect fit with the native speaker's intuitions. The same applies to the moral theorist. Should there be any deviations between an appealing account of his intuition and any moral convictions he may have previously held, then he must revise the latter until he achieves a state of equilibrium. Again, as Singer points out, this is quite in keeping with Quine's model of science. A "correct" scientific theory is one in which there is a perfect fit between theory and data. If the scientist can assure himself that he has a plausible theory on hand, then he may have to adjust his data in order to attain a perfect fit.⁽²⁴⁾

24 Ibid. p.494.

However, if these remarks are indeed correct, that is, if all that the term "validity" means for Rawls is just the conjunction of theory and first order moral judgments, then it appears that his theory is quite vulnerable to the charge of subjectivism (at least the variant of subjectivism alleged by Singer and Hare). Without some modifications, one can hardly speak of a moral theory being objectively valid from such a standpoint since the validity of any particular theory could well vary according to the one whose considered moral judgments the theory is tested against. Peter Singer attacks Rawls rather vigorously on this point. He says :

...Rawls is a subjectivist about morality in the most important sense of this often misused term. That is it follows from his views that the validity of a moral theory will vary according to whose considered moral judgments the theory is tested against. There is no sense in which we can speak of a theory being objectively valid, no matter what considered moral judgments people happen to hold. If I live in one society, and accept one set of considered moral judgments, while you live in another society and hold a quite different set, very different moral theories may be "valid" for each of us. There will then be no sense in which one of us is wrong and the other right. (25)

One way of responding to this objection is to argue in the direction of one unique reflective equilibrium for all men. If all of humanity shared the same considered judgments, then any theory that matches the considered judgments of any individual or groups will also match the considered judgments

25 Ibid.

Hare also refers to Rawls as a subjectivist for the same reasons as those expressed by Singer in this passage.

of everyone else.⁽²⁶⁾ However, Rawls does not speculate very explicitly on such a move; nor does Singer see any good deriving from it even if Rawls were to do so. Singer argues that such an appeal can, at best, only yield inter-subjective validity.

This point (the charge of subjectivism) is not affected by the question of whether there is one unique reflective equilibrium for all men or not (a question upon which Rawls refused to speculate). Even, if everyone shared the same considered judgments, this would only mean that a theory might have inter-subjective validity: it would not make for objective validity. People might have judged differently, and then a different moral theory would have been "valid". (27)

This argument is rather strange. If the considered judgment of the whole of humanity is not sufficient to make a moral judgment valid, what else can make it valid? If the validity of a moral judgment is entirely independent of the considered moral judgments of all human beings, as Singer's argument clearly implies, in what sense does one speak of a rational moral judgment? Evidently, Singer is operating with a unique sense of the term "valid" (a somewhat Cartesian sense).

26 This is of course not an easy proposition to prove. We can hardly ignore, as Steven Lukes rightly argues, the diversity of moral beliefs among men. Perhaps one could still respond to this by saying that people may yet come to agree on some common basic moral beliefs if they are properly informed and if they reason from ideal conditions. It is doubtful nonetheless, whether Rawls would actually want to proceed in this direction. His social ideal, as we will come to see, is one that allows different conceptions of the good.

27 Peter Singer, op. cit., p.494.

In at least one sense of the common usage of the term, the co-relation of a particular judgment with the considered judgments of a number of persons is an appropriate condition for accepting the validity of the former.

Whatever might be the actual merits or the demerits of postulating one reflective equilibrium for all of humanity, it seems that Rawls' notion of the reflective equilibrium can yet be shown to be plausible without such a controversial appeal. As seen in the preceding section, what Rawls wants to do in his theory of justice is to account for our sense of justice, that is, our common sense convictions and judgments about those things which we normally consider to be just or unjust, and the reasons by which we support those convictions and judgments. To accomplish this task, Rawls wants to formulate a set of principles which, when conjoined to persons' beliefs and knowledge of the circumstances, would lead them to make the judgments with their supporting reasons which they actually make. Rawls, as seen earlier, sees this task as similar to that of a person trying to construct a linguistic theory. Perhaps it should be noted here that in trying to formulate principles that regulate our sense of justice Rawls is not claiming that we actually invoke these principles when we explain or justify our particular judgments of what is just or unjust. Rather what he means is that those judgments which we do in fact make could be deductively derived from the

principles in question if these are applied to the cases in hand. In which case those judgments which we do make are actually accounted for once they could be subsumed under principles which introduce some systematic and consistent order into them.
(28)

Thus, Rawls' theory, as noted earlier, is a descriptive explanatory theory. Like a linguistic theory it appeals to "facts". The facts Rawls appeals to are the common-sense moral convictions and considered judgments which people do in fact make. And since his purpose is to explain or account for the sense of justice that regulates these judgments, it is unnecessary for him to involve himself in such questions as whether these judgments are correct, or whether or not they ought to be held or rendered. As it is in the case of the linguist, it is sufficient for Rawls to take our considered moral judgments as given. What is required of him is to delineate the standpoint from which these judgments are to be considered. And, as shown in the passage cited earlier, the

28 Aharon F. Kleinberger, *op. cit.*, p.122.
What Rawls is trying to do here is not different from what the linguist or even the physical scientist tries to do. The covering laws of science and linguistics introduce order into the observed phenomena they are said to explain. Also an instructive comparison in this kind of approach within the realm of moral philosophy is Rawls' ~~mentor~~, Kant. In formulating the "Categorical Imperative" Kant was not trying to discover what law we consciously appeal to when we act morally. Rather he wants to formulate a law or a rule that could explain those actions that we would normally describe as moral.

judgments that are to be considered relevant from the point of view of his theory are people's "considered moral judgments". These are judgments which people would make with great confidence, and which are rendered under conditions favourable for deliberation. From this standpoint, therefore, Rawls' theory of justice is the hypothesis that the principles of justice as fairness will match our considered judgments, and so describe our sense of justice better than any of the rival conceptions of justice. (29) Therefore, contrary to the viewpoint expressed by such critics as R.M. Hare and Peter Singer, the notion of reflective equilibrium may yet provide a plausible method for testing the validity of a moral theory, if the goal of the moral philosopher is to provide a descriptive explanation of our moral sense or moral capacity.

The foregoing considerations would seem to counter any wholesale condemnation of Rawls as a subjectivist simply because he appeals to people's moral judgments and opinions in testing the validity of a moral theory. As suggested earlier, if the moral philosopher cannot take the substantive judgments of actual moral agents into account, it is difficult to see what else he might appeal to in constructing a descriptive theory of morality. It does not appear that

29 Rawls, op. cit., p.48.

Rawls is doing anything different from what Hare himself has suggested in the Language of Morals. The way to find out what kind of principles govern people's moral behavior is, according to Hare, to look at what they do when they act from the moral point of view. However, there are still two thorny questions that one might still ask in connection with Rawls' descriptive-explanatory theory of justice. The first question pertains to his quest for an Archimedian point, and his claim to have found such a point in the two principles of justice as fairness. As noted earlier, in trying to explain or account for our moral sentiments (sentiments of justice), Rawls wants to formulate a set of principles from which he could deduce our actual considered moral judgments. But is there such an agreement among the considered moral judgments and convictions of different individuals regardless of such contingencies as colour, race, creed and historical periods, such that these judgments and convictions can be explained or derived from the same set of principles (the two principles of justice as fairness)?

It would appear at first that Rawls may not actually be thinking of such a wide consensus. As remarked earlier, Rawls does not speculate explicitly on a unique reflective equilibrium for all human beings. In one instance, for example, he declares that for the purpose of his book, "the opinions of the reader and author are the only ones that count."

The opinions of others are used only to clear our heads"⁽³⁰⁾.
This statement seems to indicate a very narrow consensus:
that of "one" reader of Rawls' theory and Rawls' considered
judgment himself. But, as Singer remarks, what we see here
may only be a matter of style. The very assurance with which
Rawls assumes that the reader (any random reader) shares his
own intuitions can only derive from the assumption that there
is a wide consensus about what is just and what is unjust.⁽³¹⁾
Recall an earlier remark by Rawls concerning the constraints
on knowledge at the original position. By "bracketing" all
the particularities that differentiate individuals from one
another, a choice by one person chosen at random will necessa-
rily coincide with the choice any person would make in the
original position. Similarly, if the intuitions of any rea-
der chosen at random on justice will coincide with Rawls' own
intuition, then it must be the case, if such an approach is to
be justified, that there is a wide consensus among people with
regard to these intuitions. But is this really true? Do
people as a matter of fact have the same considered judgments
about what is just and what is unjust, regardless of what
other beliefs they might have and the type of society in which
they live?

30 Rawls, A Theory of Justice, p.50.

31 Peter Singer, op. cit., p.495.

This question cannot be answered in the affirmative. As Steven Lukes rightly observes, there is a diversity of view-point among people from one society to another about what is just or unjust. The "we" and "our" to which Rawls makes frequent appeals, can only refer to a tiny segment of the human race, because, ordinarily, "we" manifestly do not agree. According to Lukes, Rawls' "we" and "our" are, in effect, no more than the relatively well-educated group who constitute generally the "liberal-social democratic consensus"⁽³²⁾. These are people who seek a world

...where there would be less inequality of wealth, no inherited privilege, no educational discrimination against social groups or the economically disadvantaged, no unearned income except on the basis of need, no inequality of reward except those based on need, merit, and contribution to the common good, equality of opportunity, no deference not based on praise, the authority of positions to be mutually agreeable in advance of their being occupied, maximum consultation before administrative decisions and unlimited comparisons between social positions in the bringing of claims against one another. (33)

The force of Lukes' criticism will become evident when we come to review Rawls' index of primary goods. Although Rawls would have us believe that the primary goods are just what people will seek if they are rational, they are, in actual fact, no

32 Steven Lukes, "Relativist Cognitive and Morals". Arist. Society Supp., 48, (1974), p.182.

33 Ibid.

more than the things typically sought by Western liberal democrats.

Now, if Lukes is correct, that is, if there is no common consensus among people in general on considered judgments about justice; and if it is indeed the case that the considered judgments which Rawls has in mind (the ones he is offering for the test of justice as fairness) only reflect the values of Western liberal democrats, does it not follow then that the validity of his theory can only be relative to the moral opinions of a particular group of people? Does this not further indicate the need if we are to attain a genuine Archimedian point, for various sets of principles or theories of justice to account for all the disagreements in people's considered judgments about what counts as just?

Within the present Western societies, there are, as Lukes points out, "ultra-conservatives, clerical authoritarians, Empire loyalists, individualist liberals, anarchists, radical egalitarians"⁽³⁴⁾. Evidently, the considered judgments of anyone in these societies would differ essentially from those of Rawls' liberal men. And if this is the case, how can the reflective equilibrium reached by anyone in the latter societies coincide with the one reached by Rawls' men? Should they not coincide, why must we prefer one to the other?

34 Ibid.

Perhaps one might respond here that Rawls' considered judgments should be accepted because they are "the right ones, the most rational, the correct ones, the ones reasonable men should accept".⁽³⁵⁾ It might be possible to argue in this direction. But Rawls has not done so, nor is it certain whether he can actually do so within the framework of his theory. If he intends his theory to be descriptive-explanatory, then he cannot at the same time be preoccupied with the question of whether the actual judgments and convictions he wants to account for are correct, or whether or not they ought to be held. He must accept these judgments as given and then proceed to account for them.

Rawls seems to anticipate this objection. For instance he admits rather explicitly that he does not have a conclusive answer to the challenge of pluralistic intuitionism. However, it is his belief that at least one educated person's (who happens to be himself) views about justice do coincide with the principles of justice as fairness. This is obviously true. But how such a 'one' man agreement could yield an Archimedian point remains very unclear.

The second question raised by Rawls' descriptive-explanatory hypothesis pertains to the relevance of his device

35 Ibid.

of the original contract. If the two principles of justice as fairness are to be considered as a descriptive-explanatory hypothesis which must be tested against people's actual moral judgments and convictions in order to determine the validity of these principles, how does the justificatory role of the original position enter into the picture? It is not the entire construction of the original position that is being called to question here. From the standpoint of a descriptive-explanatory theory of justice, Rawls might still need an original position as a way of discovering what principles actual moral agents would in fact endorse, and also, what motivational considerations would lead them to endorse such principles. But this is only a question of discovery and not a question of justification. The analytic role of the original position as expounded in Chapter One fits into this descriptive role. Also, as Kleinberger points out, the contractarian strategy in a descriptive role would be a "way in which one may light upon a fruitful and promising hypothesis"⁽³⁶⁾. The question is that of psychology and not logic. But the question of justification, by contrast, is essentially that of logic. Within the scientific framework it involves the testing of the truth or validity of an hypothesis by its observational consequences, regardless of how the hypothesis has been generated.⁽³⁷⁾ The primary concern

36 Aharan Kleinberger, op. cit., p.123.

37 Ibid. p.124.

here is not what principles would form the hypothesis but why the principles contained in the hypothesis must be accepted. The social contract does not have any useful role to play in such a question since its main concern is to find out what principles would best describe a well-ordered society of moral persons.

Rawls however, wants his contractarian account of justice to be at one and the same time descriptive-explanatory as well as prescriptive-justificatory. His remarks about what kind of theory a theory of justice and morality ought to be, indicate a move in the direction of a descriptive-explanatory hypothesis. Here the analogy between the moral philosopher and the linguistic theorist is very crucial. Both theorists as it were are pre-occupied with an attempt to formulate principles or theories that would correctly explain first order linguistic intuitions of the native speaker (in the case of the linguist), and first order moral intuitions of moral persons (in the case of the moral philosopher). The first order intuitions in both cases are taken for granted or considered incorrigible. But in the statement of the reflective equilibrium, as cited earlier, Rawls does not consider the actual judgments and convictions of moral persons as incorrigible. The principles of justice are to prescribe what opinions and judgments persons may have in particular instances. Should there even be a conflict between one's existing judgments and an intuitively appealing

account of one's sense of justice, then the former ought to be adjusted to fit the latter.

When a person is presented with an intuitively appealing account of his sense of justice (one, say, which embodies various reasonable and natural presumptions), he may well revise his judgments to conform to its principles even though the theory does not fit his existing judgments exactly. He is especially likely to do this if he can find an explanation for the deviations which undermines his confidence in his original judgments and if the conception presented yields a judgment which he finds he can now accept.

(38)

It seems rather obvious from this remark that the validity of a moral theory does not rest merely in trying to capture and account for the actual considered judgments that people do in fact make. Rather, a particular theory of justice is considered justified only when its principles correctly prescribe what our moral opinions and judgments ought to be. (39)

The primary goal of Rawls' contractarian strategy is to achieve the latter type of justification. The choice of the two principles of justice as fairness in the original position is to constitute the justification of these principles. By appealing to these principles Rawls wants to provide a strategy for assessing the justness of basic institutions and for judging all moralities. However, from the argument

38 Rawls, op. cit., p.48

39 And how is this finally ascertained? Here is where the problems or approach of intuitionism seem to remain.

in the previous chapter, it seems evident that the design of Rawls' original position does not guarantee the emergence of principles which must be accepted from the moral point of view.

One of the fundamental errors in Rawls' moral methodology is his failure to distinguish between the two types of theories he was trying to present. Rawls tries to conflate both types of theories into one by means of his notion of reflective equilibrium. To attain the state of equilibrium an individual would shuttle between the principles of justice and the underlying account of the initial contract situation, on the one hand, and his considered convictions of justice, on the other. One is at liberty to alter the conditions of the contractual circumstances, and if the need arises, one may even withdraw one's previously held considered judgments or try to fit them into a more satisfactory conception of justice. Rawls of course fails to tell us what specific criteria must control these shuttlings and revisions. The individual is only required to shuttle and revise his judgments in order to bring them in line with "an intuitively appealing account of his sense of justice"⁽⁴⁰⁾. But then Rawls is yet to tell whether such an intuitively appealing account of a person's sense of justice will necessarily coincide with that of someone else; and if it does, why it does.

40 Rawls, A Theory of Justice, op. cit., p.48.

Thus, it would seem that, whereas the very idea of appealing to the considered moral judgments of actual agents in testing a moral theory, is by itself a plausible strategy, Rawls' failure to distinguish the proper context in which such a method can be used has diminished its effectiveness. In fact, the failure to articulate this distinction has led to some very important confusions about the entire argument of the reflective equilibrium, especially how it relates to other arguments in justice as fairness. There are three arguments characterizing the notion of justice as fairness (contract, equilibrium and congruence). Although each of these arguments deals with specific aspects of the theory, it is doubtful whether Rawls actually intends any of them as a sufficient condition for the validity of his theory. Rather, it would seem that each of the three arguments is intended as a mutual support for the other: the reflective equilibrium as a support for the contract argument, and the congruence as a support for both the reflective equilibrium and congruence arguments. Rawls says quite explicitly that a moral theory that does not take into account the consequences that might follow from it, is irrational and crazy. Hence, one of the conditions of validity for a moral theory is the probable consequences that might result from it: its congruence with human good. And to show that justice as fairness is congruent with human good, Rawls initiates the congruence argument which

occupies the entire third section of A Theory of Justice. Thus, one of the features that must be considered in assessing the moral methodology of Rawls is the congruence argument. This is the final justification of justice as fairness. This feature will be examined in Part Two of this study.

Meanwhile, one might conclude that some of the serious criticisms which Rawls' moral methodology has drawn, arise mainly because of the unclarity of the status of his theory. Rawls has not been very helpful in helping his readers determine what model of theory he is developing in a theory of justice. This general survey of the main features of Rawls' theory will conclude with an examination of the two principles of justice as fairness.

CHAPTER THREE

THE PRINCIPLES OF JUSTICE AS FAIRNESS

Hitherto, the focus of this survey has been the theoretical considerations which Rawls advances in support of his principles of justice. The goal has been to assess the validity of Rawls' methodological procedure. In this chapter we take a look at the principles themselves. We recall from our earlier discussion that Rawls considers his two principles of justice established, if they would be chosen by the parties in the original position, and, if they would conform with our ordinary considered judgments duly pruned and adjusted.⁽¹⁾ Now to what extent have these two conditions been met? Emphasis in the following discussion will focus on the second condition: Is it the case that the two principles of justice as fairness do in fact illuminate and make precise most people's considered judgments about justice? Also, do these principles help reveal in anyway a basic structure and coherence underlying our ordinary intuitions about justice? We begin our examination with the first principle of greater equal liberty.

1 Rawls, A Theory of Justice, p.20

THE PRINCIPLE OF GREATEST EQUAL LIBERTY

RESTRICTION OF LIBERTY FOR THE SAKE OF LIBERTY:

All through his book, Rawls emphasizes a distinction between liberty and other primary social goods. His special conception of justice is very explicit about this distinction. Unlike the general conception of justice, the special conception embodies a priority rule which assigns to liberty a priority which forbids its (liberty) restriction except for the sake of liberty:

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.... Liberty can be restricted only for the sake of liberty. There are two cases: (a) a less extensive liberty must strengthen the total system of liberty shared by all; (b) a less than equal liberty must be acceptable to those with the lesser liberty. (2)

Rawls intends the special conception to govern societies whose evolution has reached such a point that, according to him, "the basic wants of individuals can be fulfilled", (3) and social conditions permit "the effective establishment of fundamental rights". (4) In the absence of such favourable conditions, then

2 Ibid. p.302.

3 Ibid. p.543.

4 Ibid. p.542.

equal liberty may be denied, if this would help "raise the level of civilization so that in due course these freedoms can be enjoyed".⁽⁵⁾ Liberty is only to be restricted for the sake of liberty itself.

To enter into a deeper meaning of this complex doctrine, it might be necessary to take into account not only what Rawls says expressly about liberty, but also what he says about such notions as individual autonomy, natural duties, obligations arising from the principle of fairness and the common good or common interest. These notions -- in particular, individual autonomy will help us see more clearly both the background from which the principle of greatest equal liberty emerges, and how it differs in scope from its earlier version: the general principle of liberty in Rawls' earlier papers.⁽⁶⁾ A main reason -- to cite an important case in point -- why the contract approach appeals so strongly to Rawls stems from his belief that the constraints of right and fairness require that people should submit to procedures and institutions only if, given the opportunity, they would in some sense have agreed in advance on the principles to which they must submit. He declares that "a society satisfying the principles of justice as fairness comes as close as a society can to being a voluntary scheme, for it meets the principles

5 Ibid. p.152.

6 John Rawls, "Justice as Fairness", Philosophical Review Vol.67, 1958, 164-94.

which free and equal persons would assent to under circumstances that are fair".⁽⁷⁾ Thus, as Thomas Nagel rightly observes, the "fundamental attitude toward persons on which justice as fairness depends is a respect for their autonomy or Freedom."⁽⁸⁾ The precise task of the principle of the greatest equal liberty is to capture and preserve this autonomy.

One other initial observation that may yet be made is the apparent similarity between Rawls and Herbert Spencer on the notion of liberty. In his Social Statics, Spencer conceives the law of equal liberty as an elementary truth of ethics. "Every man", he says, "has freedom to do all that he wills, provided he infringes not the equal freedom of any other man".⁽⁹⁾ This assertion seems to translate adequately the general idea in Rawls' claim that liberty can only be restricted for the sake of liberty itself.

However, Spencer's principle has been vigorously criticized by Sidgwick in the Method of Ethics. Sidgwick argues that the principle fails to account for some of the most obvious restrictions on liberty, which are required not to promote greater liberties but to protect individuals from harm and to secure other material benefits. For example, several societies, as

7 John Rawls, A Theory of Justice, p.13.

8 Thomas Nagel, "Rawls on Justice", Reading Rawls, p.4.

9 Herbert Spencer, Social Statics, p.56.

Sidgwick points out, forbid the institution of private property in order to safeguard the public interest: Having the right to own a thing privately is to have the liberty to use that thing in ways denied to others. (10)

Rawls appears to have taken Sidgwick's criticism of Spencer into account in A Theory of Justice. While his general principle of equal liberty as it appears in his earlier writings may indeed be similar to the one criticized by Sidgwick, his special formulation of the same principle in his book tries to avoid these difficulties. In the latter formulation, Rawls refers not simply to "liberty" but to "basic liberties". These basic liberties, as H.L.Hart points out, are the sort of things both legally recognized and protected from interference in our society. (11) Thus Rawls' revision of the equality of liberty principle in A Theory of Justice would seem to render the principle immune to the type of criticism which Sidgwick raises against Spencer's formulation of it. (12) Perhaps what we need to do now is to probe the claim that "basic" liberties can only

10 Henry Sidgwick, The Methods of Ethics, 7th edition, pp.275-6; pp.297-9.

11 H.L.A.Hart, "Rawls on Liberty and its Priority" Reading Rawls, p.235.

12 While this revision may indeed enable Rawls to avoid the sort of criticism raised by Sidgwick, it does raise some questions of its own. For instance, does the change in the formulation of the liberty principle in Rawls' book indicate a change in view point? Or, is the principle of liberty in

be restricted for the sake of greater basic liberties. The question to ask here is rather straightforward: do reflective persons only restrict basic liberties for the sake of more basic liberties? In other words, is it always wrong to restrict basic liberties for any other reason apart from increasing basic liberties?

To discuss this question effectively, it is necessary to have a precise representation of what the claim itself is. What does the idea of restricting "basic liberties for greater extent of basic liberties" convey? What are basic liberties, and how does one distinguish them from other kinds of liberties? Rawls does not provide any substantive definition of what he means by "basic liberty". Rather he suggests a range of what may be regarded as basic liberties with a number of examples. These examples include (a) the restriction of individuals' liberty to interrupt communication in a debate (the liberty to speak just when they want without rules governing when they may speak), (b) the restriction of the freedom of individuals through

A Theory of Justice still the same as the general principle, in which case, under the priority rule attached to it, no form of liberty may be restricted except for the sake of liberty? There are several indications in Rawls' book that one might interpret as a real change in theory. For instance, Rawls does not go into the trouble of attempting to reconcile the admission of private property as a liberty with any general principle of maximum equal liberty. Also, by making a distinction between liberty and the value or worth of liberty (a distinction which Norman Daniels effectively criticizes as being arbitrary - see "Equal Liberty and Unequal Worth of Liberty" in Reading Rawls, pp.259-263), Rawls tries to provide a new meaning to the requirement that the right to hold

military conscription in times of just wars undertaken in defence of free institutions, and (c) the restriction of the right to shout out at will in such public places as theatres, auditoriums, churches and the like and thereby interrupting a cinema, a play, lecture or church service. While these examples do appropriately illustrate Rawls' principle, his interpretation of them, as H.L.A. Hart and Kai Nielsen point out, are quite misleading. While people generally would agree that such restrictions are indeed necessary, it is doubtful whether the restrictions would be seen by every reflective person as an attempt to increase greater basic liberties. Let us consider the example of the restriction of freedom of speech in a debate.

Rawls cites the introduction of rules of order in a debate as a typical example of restricting basic liberties for the sake of greater basic liberties. He argues that while the

property must be equal thereby eliminating the problem which Spencer ran into. Rawls also insists repeatedly that while he recognizes the right to private property as a liberty, the choice between capitalism and state ownership of the means of production is left quite open by the principles of justice (see pp.66, 273-4). If these observations are correct, then one might say that the reformulation of the liberty principle is not just a change in language but a change in perception. Unlike the general principle, it focuses entirely on basic liberties. Now, should this indeed be the case, is one then to expect another principle that would regulate the less than basic liberties? It is not immediately clear how Rawls would want to respond to this question, nor is it evident that the second principle (the difference) is specifically designed to deal with the latter category of liberties.

basic liberty of individuals to speak just when they want may appear restricted by the rules of order in a debate, the restriction aims at, and actually procures greater liberty for individuals to say and advocate what they want. The aim of the restriction, he says, is to gain the very freedom of speech it restricts:

...Certain rules of order are necessary for intelligent and profitable discussion. Without the acceptance of reasonable procedures of inquiry and debate, freedom of speech loses its value.... While rules of order limit our freedom, since we cannot speak whenever we please, they are required to gain the benefits of this liberty. Thus the delegates to a constitutional convention, or members of the legislature, must decide how the various liberties are to be specified so as to yield the best total system of equal liberty. They have to balance one liberty against another. The best arrangement of the several liberties depends upon the totality of limitations to which they are subjected, upon how they hang together in the whole scheme by which they are defined. (14)

This example indicates two conflicting liberties: the liberty to communicate one's thought in speech and the liberty to interrupt communication. Even though the latter might seem rather negative, there are instances in which an individual can rationally exercise his autonomy by way of interrupting communication. (Recall the parliamentary practice of interrupting a

13 H.L.A. Hart, op. cit., Kai Nielsen, "The Priority of Liberty Examined", The Indian Political Science Review, vol. xi, Jan., 1977, pp. 48-59.

14 Rawls, A Theory of Justice, p.303.

speaker "on point of order"). However, when one takes into account the very purpose of having a debate (to thrash out issues of mutual concern) then it becomes rather obvious that the liberty to communicate one's thought orderly in speech is more profitable and hence desirable. Thus Rawls is quite correct in suggesting that the liberty to interrupt communication can be legitimately restricted in favour of the liberty to communicate in speech. However, as Hart points out, there is something yet misleading in the way Rawls stresses the notion of liberty in this rather straight forward example :

It seems.... misleading to describe even the resolution of the conflicting liberties in this very simple case as yielding a "greater" or "stronger" total system of liberty for these phrases suggest that no values other than liberty, and dimensions of it, like extent, size, or strength, are involved. Plainly what such rules of debate help secure is not a greater or more extensive liberty, but a liberty to do something which is more valuable for any rational person than the activities forbidden by the rules, or, as Rawls himself says, something more "profitable". So some criterion of the value of different liberties must be involved in the resolution of conflicts between them; yet Rawls speaks as if the system "of basic liberties" were self-contained and conflicts within it were adjusted without appeal to any other value besides liberty and its extent. (15)

Hart's argument is quite convincing. As mentioned above, any thoughtful person will, on reflection, agree that the liberty to communicate one's thought in speech in a debate should override the liberty to interrupt communication. But the basis of

15 Hart, op. cit. pp.239-240.

such an agreement is not always to enhance or secure greater liberties but to pursue the common objectives which led to the debate in the first place. In fact unless one is appealing to the "greatest happiness to the greatest number" principle it is difficult to see how the restriction on liberty here could be shown as aiming at a greater extent of liberties for everyone. For one thing there may be individuals who only want to exercise the liberty to interrupt communication. The liberties of such individuals can hardly be said to be increased, if the only liberty they choose to exercise is restricted. Perhaps we might want to say that if such individuals are rational then they ought to seek the liberty to communicate in speech and not exercise the liberty to interrupt communication. But even then we still need to take into account what these people want as a matter of fact. If interrupting communication is quite consistent with the end they seek, then we have no basis, at least not from Rawls' point of view to regard them as irrational. Thus, Hart seems to be perfectly justified in indicating that even in such a relatively simple example as the rules of order in a debate, it is not so obvious that the purpose of restricting the freedom of speech is simply to regain the same freedom. These rules aim at securing liberty not as an end itself but as a means to do whatever the debaters want to accomplish (that which they consider more valuable than the liberty to disrupt communication). We

do not arrive at such restrictions by a mere intuitive comparison of one liberty with another. Rather what we do is to balance the benefits of each liberty against the benefits of the other. And, a decision is made in favour of the liberty whose benefits are more in keeping with the overall objective of the debate. The reasoning process is essentially utilitarian. Rawls himself seems to be well aware of this fact especially when he resorts to such phrases as "profitable discussion", "benefits of this liberty" etc., in the statement of his example.

Rawls' claim (that liberty can only be restricted for the sake of liberty) becomes even more vulnerable when we consider conflicts that are more complex than the one suggested in the debate example. As mentioned above, people would generally agree on reflection that the liberty to interrupt communication in a debate should be restricted in favour of the liberty to communicate our thought in speech. But should we consider a more complex case, say, one involving conflicting liberties in a constitutional arrangement, it is not so evident that a consensus can easily be reached. Suppose, as Nielsen remarks, we consider the conflict concerning liberties implicit in two conflicting proposals in designing institutions. One proposal allows for unlimited majority rule, and the other restricts majority rule through some constitutional device. Rawls of course opts for constitutional limi-

tation on majority rule in the hope that this option can be effectively justified by an appeal to greater equal liberty. But this is not so evidently the case. Rawls himself admits that reasonable persons may yet mutually disagree with one another about the value of the conflicting liberties in a situation as this, and this in turn will affect their perception of the conflict itself. For instance, classical liberals would be inclined to attach more intrinsic importance to liberty of conscience and freedom of the person than political liberties. As such, they could, in a choice situation, opt for a sovereign rule (provided the latter guarantees the rule of law, liberty of conscience and freedom of person to unlimited majority rule). But more progressive liberals such as Rawls himself would hardly opt for this. They would instead prefer political liberties even if this in turn meant a lesser extent of liberty of conscience. Now how do we resolve such a conflict by simply appealing to greater equal liberties? How does one tell which of the two options would lead to greater equal liberties without appealing to some other consideration. Rawls of course wants us to assess "the relative total importance of the different liberties". But "total importance" as regard to what? To conduct such an assessment, as Nielsen points out, an appeal must be made

16 Kai Nielsen, op. cit., p.52.

to other considerations apart from considerations of liberty and some of these would seem at least to be to utilitarian considerations. (17) But if this is indeed our final resort, does it not raise a question about the priority of liberty principle? If there is no other means of resolving conflicting basic liberties except to appeal to some utilitarian consideration (value or benefits of basic liberties) then it is not just plainly true that basic liberties can only be restricted for greater equal basic liberties.

Rawls does have a response to this objection. He notes that such complex situations of conflicting liberties are not common place. And when they do arise, it should be possible to find a constitutional procedure that would protect both the principle of participation as well as the other liberties. We might apply what he refers to as "the principle of mutual advantage".

Fortunately however, we do not often have to assess the relative total importance of the different liberties. Usually the way to proceed is to apply the principle of equal advantage in adjusting the complete system of freedom. We are not called upon either to abandon the principle of participation entirely or to allow it unlimited sway. Instead, we should narrow or widen its extent up to the point where the danger to liberty from the marginal loss in control over those holding political power just balances the security of liberty gained by the greater use of constitutional devices. The decision is not at all or nothing affair. It is a question of weighing against one another small variations in the extent and definition of the different liberties. The priority of liberty does not exclude marginal exchanges within the system of freedom. Moreover, it allows although it does not

require that some liberties, say those covered by the principle of participation, are less essential in that their main role is to protect the remaining freedoms. Different opinions about the value of liberties will, of course, affect how different persons think the full scheme of freedom should be arranged. Those who place a higher worth on the principle of participation will be prepared to take greater risks with the freedoms of the person, say, in order to give political liberty a larger place. Ideally these conflicts will not occur and it should be possible, under favourable conditions anyway, to find a constitutional procedure that allows a sufficient scope for the value of participation without jeopardizing the other liberties. (18)

Apart from the fact that Rawls is rather vague on what this "principle of equal advantage" is and the considerations underlying it, this response still leaves intact the objection which Hart and Nielsen have so well articulated. How can one conceivably conduct such an exercise of weighing and balancing the value of the conflicting liberties if the final appeal is liberty itself? Consider two types of democratic institutions A and B. Institution A allows for greater liberty of conscience while institution B allows unlimited majority rule. Now how does a person tell which of these two institutions makes for greater total liberty? Rawls insists that the priority of liberty principle allows trade-offs in basic liberties. This is fair enough. Such trade-

17 Ibid.

18 Rawls, A Theory of Justice, p.230.

offs are not uncommon in everyday life. But precisely how does one do this if the one and only consideration is to increase the greater extent of liberties? If the exponents of institutions A (the classical liberals who place a higher premium on the liberty of conscience and the freedom of the person as against political liberty) should accept any trade-off, there does not seem to be a sure way in which these same people can end up with a greater extent of basic liberties. The same applies to the defenders of institutions B. Indeed, should both groups make some trade-off in liberties, they might be able to evolve a system that would better fulfill their collective interests and needs. But this is not the issue. For Rawls, the important thing is to see which of the two institutions would make for a more extensive system of liberties. It does not seem that there is a sure way to know this, if all we have for a frame of reference is liberty itself. Rawls' note of optimism at the end of the text does not throw any new light on the situation. Saying that "ideally these conflicts will not occur" is not to tell us how we are to resolve them if and when they do occur. Also, his earlier remark that when such conflicts occur, we could find out what a representative equal citizen would find it rational to prefer is equally unconvincing. Apart from the fact that the latter is a standard utilitarian procedure (a procedure that Rawls explicitly rejects), such an appeal would only work if there is a sure way of defining the common

interests of the parties. This might be possible with regard to the restrictions imposed by the rule of debate. Those who would want to press for the liberty to interrupt communication could be shown to be irrational by merely appealing to what a representative person would, on reflection, consider to be the goal and purpose of a debate.

Here a consensus might easily be reached. But in the latter case of the constitutional arrangements, a consensus is not so easily attainable. Besides, as Hart rightly observes, "it would be quite wrong to generalize from this single case (debate); other conflicts between basic liberties will be such that different resolutions of the conflict will correspond to the interests of different people who will diverge over the relative value they set on the conflicting liberties. In such cases, there will be no resolution which will be uniquely selected by reference to the common good". (19)

Thus Rawls' principle of the priority of liberty seems to encounter serious difficulty from the very onset. Unless an appeal is made to some utilitarian consideration, one can hardly avoid arbitrariness in the restrictions of basic liberties. It is almost impossible to know the representative person's rational preferences; nor can one tell how such preferences might yield a greater system of liberties

19 Hart, op. cit., P. 241

It is true, as was mentioned earlier, that reasonable people do make trade-offs in liberties in everyday life. While some of these trade-offs are made with a view to obtaining greater liberties, liberties are also restricted in a direct attempt to procure material benefits. In legal proceedings, a person might waive some of his rights in return for more extensive rights. But as Hart has convincingly argued, liberties are also waived or restricted to promote such material benefits as good health, prevention of bodily suffering and economic well being.

Rawls' uncritical acceptance of certain ideals of liberalism does not allow him much room to manoeuvre on this point. For instance, as mentioned earlier, his acceptance of the notion of contract as the appropriate model for a theory of social justice, and in fact, as also a possible basis for constructing a theory of justice between individuals rests upon the assumption that it is only fair to require autonomous individuals to submit to procedures and institutions only when they must have previously agreed (in some sense) on what principles to adopt. Thus the notion of individual autonomy is very crucial in his thought experiment.

Individuals can be said to have made a rational choice only if the principle on which they chose is autonomous. For him (Rawls), as it is for Kant, "... a person is acting autonomously when the principles of his action are chosen by him as the

most adequate possible expression of his nature as a free and equal rational being. The principles he acts upon are not adopted because of his social position or natural endowments, or in view of the particular kind of society in which he lives or the specific things that he happens to want. To act on such principles is to act heteronomously.⁽²⁰⁾ Thus by imposing the veil of ignorance, Rawls believes that he has successfully deprived "the persons in the original position of the knowledge that would make them choose heteronomous principles"⁽²¹⁾.

Perhaps individual autonomy is indeed crucial to the evolution of principles of justice and fair play. But, given Rawls' specification of autonomy above, does any man ever act autonomously? The real world as it exists is one in which individuals, while cherishing their autonomy, are nevertheless, prepared to restrict it in the interest of other values as common sense and circumstance dictate. It seems rather arbitrary for Rawls to insist that liberty or autonomy must always override all other values: (at least he has not argued for this). In the examples of the rules of debate mentioned earlier, while it is indeed important for us as moral beings that where important issues are under consideration, we want

20 Rawls, A Theory of Justice, p.252.

21 Ibid.

the debate to be fair, the fairness of the debate is not an end in itself but a means toward the resolution of the issues being considered. Hence common sense as well as the nature of the issues may constrain the parties to, on some occasions, override the factor of individual autonomy and proceed to a constructive debate along what Rawls might call heteronomous approach. Both approaches could be quite consistent with our ordinary considered judgments depending on the prevailing circumstances.

A father of a family who waives some of his liberties in the interest of peace in his family need not be seen to be doing so for greater liberties. And, there is no reason to regard such a restriction as irrational just because it does not aim at greater liberties. Thus, contrary to Rawls' claim, the principle of the priority of liberty may not make an equal appeal to every rational person as being uniquely rational. We now turn to the difference principle.

SECTION TWO

THE MAXIMIN CRITERION

The second principle of justice as fairness is directed specifically to other "primary goods" apart from liberty and self-respect, to the extent that the distribution of these primary goods is affected by the basic structure of social institutions. Since Rawls expects us to reason from the standpoint of a well-ordered society of free and equal

moral persons, we might naturally suppose that all the social primary goods, especially income and wealth ought to be equally distributed: that is, every person should receive an equal share of all other social primary goods along with basic liberty and self-respect. An equal distribution of these assets would seem to be much more in keeping with the equality of free moral persons. However, as Rawls points out, we need to go beyond the equality of liberty principle in view of the organizational requirements and the constraints of economic efficiency in the society. Hence there is a need for a criterion of distribution that allows certain inequalities:

...Society must take organizational requirements and economic efficiency into account. So it is unreasonable to stop at equal division. The basic structure should allow inequalities so long as these improve everyone's situation, including that of the least advantaged, provided these inequalities are consistent with equal liberty and fair opportunity. Because we start from equal shares, those who benefit least have, so to speak, a veto; and thus we arrive at the difference principle. (22)

Thus, concerning such primary goods as income, wealth, power and authority, the second principle of justice as fairness directs that "social and economic inequalities are to be arranged so that they are both (a) to the greatest benefit of the least advantaged and (b) attached to offices and positions

22 John Rawls, "A Kantian Conception of Equality", Cambridge Review, February 1975.

(23)
open to all under conditions of fair equality of opportunity. Although the second part of this principle may be classified as an independent criterion, it shall be ignored hereafter since its relevance is essentially duplicated by the greatest equal liberty principle. The emphasis, rather, will be on the first part which Rawls refers to as the "difference principle". Before considering any serious objections to the principle, it might be useful to illustrate how Rawls conceives its superiority over the principle of utility.

There are, according to Rawls, several initial (prima-facie) reasons that support the choice of the difference principle (the maximin criterion) over the principle of utility in the original position. First, under the special features of the original position, the choice of the maximin criterion does not demand excessive risk on the part of the contractors: the information requirements for its application is far less than that required for the application of the utility principle.

...much less information is needed to apply the maximin criterion. Once the least-favored group is identified, it may be relatively easy to determine which policies are to their advantage. By comparison it is much more difficult to know what maximizes average utility. We require a fairly precise way of comparing the utilities of different social groups by some meaningful standard, as well as a method of estimating the overall balance of gains and losses. In

application this principle leaves so much to judgment that some may reasonably claim that the gains of one group outweigh the losses of another while others may equally reasonably deny it. This situation gives those favored by existing inequalities an opportunity to exploit their advantage so that, as a result, inequalities are likely to be excessive, undermining the justice of the system. (24)

The reasoning in this passage does not seem very decisive as Rawls himself subsequently admits. If it is assumed, as he does, that a person's utility is affected predominantly by such contingencies as liberties and opportunities, income and wealth, and if it is the case that everyone has normal physical needs, it should be possible to compare the utilities of different social groups. An appeal to the primary goods would seem to be equally effective within a utilitarian framework for the purpose of assessing the overall balance of gains and losses. Perhaps the exercise may indeed be more cumbersome in the case of finding out what maximizes average utility. But this at best would be a practical and not a theoretical difficulty. It should not make the utility principle any less valid than the maximin criterion.

24 John Rawls, "Some Reasons for the Maximin Criterion", American Economic Ass. May 1974, pp.143-4.

25 Rawls' definition of the utility principle might seem to counter this claim. As he says, the principle of average utility is "to be taken in the sense, which permits (by definition) inter-personal comparisons of utility that can at least be assessed at the margin; and utility is to be measured from the standpoint of individuals in society (and not from the standpoint of the original position) and means the degree of satisfaction of their interests". (Rawls,

Another initial consideration that Rawls advances in favour of the maximin criterion is that it enjoys greater suitability as a public principle in comparison to either "pattern criteria" or the utility principle.

...a distributive criterion is to serve as a public principle. Citizens generally should be able to understand it and have some confidence that it is realizable. Pattern criteria, those that require the actual distribution to exhibit certain ascertainable features, do well by the test of publicity. Of these, strict equality (equal division) is the sharpest principle. The trouble with pattern criteria is that sharpness is not the only consideration, and they often have little else to commend them. On the other hand, the utility principle is not sharp enough: even if it were satisfied, there could be little public confidence that it is indeed the case. The maximin criterion has a sufficient sharpness; at the same time, it is efficient while strict equality is not. (26)

There can hardly be any objection to Rawls' insistence on the constraint of publicity with regard to the principle of distribution. Before the members of a society can come to have a general and an effective desire to govern themselves by any given principle, it is necessary that they know what the principle requires of them and that everyone, or, at least,

"Reply to Alexander & Musgrove", p.640.) It is not very clear at this point why such an index as that of the primary goods cannot be applied directly to individuals in society. As mentioned earlier, there is usually a tight link between the wealth an individual has and the liberties or authority he exercises in society. This point will be more carefully examined later in this chapter.

26. Rawls, op. cit., p.144.

a majority of them would abide by its constraints. Nevertheless, the reason advanced by Rawls in this passage does not seem any more decisive than the one considered earlier. He merely asserts without argument that the utility principle is not sharp enough and that the public confidence in it can only be minimal. Is this so plainly true? Is it the case that any reflective group of persons regardless of their particular sets of beliefs (social, political and moral) will just not have confidence in the utility principle if it is to be a rational group? Perhaps this might be the case in actual fact. But apart from merely stating that it is, Rawls has not shown us why it is, or why it ought to be the case or even that it is.

The third ground supporting the maximin criterion is the weaker strains of commitment it entails. In contrast to the utility principle, the maximin criterion does not, according to Rawls, make an extreme psychological demand: its compliance is much more tolerable than the utility principle:

...in the original position the parties are to favor those principles compliance with which should prove more tolerable, whatever their situation in society turns out to be. The notion of a contract implies that one cannot enter into an agreement that one will be unable to keep. By this test, also, maximin seems superior, for the principles chosen would regulate social and economic inequalities in the basic structure of society that affect people's life-prospects. These are peculiarly deep and pervasive inequalities and often difficult to accept. Looking first at the situation of the less advantaged, the utility principle asks them to view the greater advantages of others who have more as a sufficient reason for having still

lower prospects of life than otherwise they could be allowed. This is an extreme demand psychologically; by contrast, the maximin criterion assures the less favored that inequalities work to their advantage. (27)

Again, the constraint of commitment seems quite reasonable in itself if individuals must choose the type of principles that are to govern their social interactions. It hardly makes sense for individuals to come together to select a set of principle if none of them has the intention to abide by such principles. Also, for rational individuals to give their consent to be bound by any principle, they must have some assurance that the demands of that principle will not be too extreme. Now it seems true, as Rawls points out, that the emphasis of the utility principle tends to be on the greater advantages that a greater number of citizens can derive from it. The fact that there might be some minority who may be constrained to work for the benefits of the majority, would tend to make the parties in the original position hesitant in choosing the utility principle. Since they are presumed to be free and equal, but mutually disinterested in one another, it will be difficult to persuade any one of them to accept less benefits in order to promote the good of the majority.

However, the maximin criterion also runs into the same difficulty. As Rawls points out, those who are better situ-

ated would be required to receive less benefits under the maximin criterion than they would otherwise receive under the utility principle. But in spite of this fact, Rawls still thinks that the strains of commitment are yet lessened considerably under the maximin criterion. The more advantaged parties are, after all, more fortunate and enjoy the benefits of the fact; and insofar as they value their situation relatively in comparison with others, they give up that much less. In fact, our tendency to evaluate our circumstances in relation to the circumstances of others suggests that society should be arranged so that if possible all its members can with reason be happy with their situation. (28) This explanation cannot be adequate if we take into account the motivational assumptions in the original position. The parties are only concerned to further their own individual interests. They do not compare themselves with others; nor do they envy one another. Now with such a motivational assumption it is difficult to see how the more advantaged ones among them would be content with accepting less benefits just because some of them are in far worse situations. It is true that in real life, one tends to overlook minor misfortunes in the knowledge that one is still better off than the average citizen despite such misfortunes. A story was once told of a man who went

28 Ibid, p.144.

on the street swearing at God for creating him with one lamed limb. But as soon as he sees someone else who is both lamed and blind, he forgets his own plight and begins to praise God. This story seems to capture the rationale of the maximin criterion. Although people generally would be unhappy to receive less benefits from their abilities, they might yet be happy when they come to realize that they are still better off than the average citizen in spite of the fact that they received less benefits. However, the two situations are not actually the same. The change in attitude of the lamed man is not due to the fact that he can still adjust to, and ultimately enjoy his situation. Rather, he became accommodating because he had balanced his own situation against that of the other person, and when he saw that his was much preferable, only then did he become relieved. The same man would certainly be very bitter all over again, the next time he goes on the street and finds out that everyone else was in a better physical condition.

Now such inter-personal comparison is just something that Rawls denies his men at the original position. None of them would derive his satisfaction from the fact that he is well off than someone else. The only thing that would make them happy is the awareness that each one is able to maximize his expectable utility, regardless of how anyone else fares. But the more advantaged can hardly consider their utility maximized if they are to accept lesser benefits than they would otherwise receive from their abilities.

Thus the superiority of the maximin criterion over the utility principle in regard to the strains of commitment is at best contestable.

It turns out therefore, that none of the three reasons advanced by Rawls in support of the maximin criterion is clearly decisive. But Rawls seems well aware of this. Hence, for him, and perhaps for everyone else, "the question arises whether there is any consideration that is compelling".⁽²⁹⁾

In answer to this question, Rawls maintains that "the aspirations of free and equal personality point directly to the maximin criterion".⁽³⁰⁾

Rawls admits that some genuine form of democracy may yet obtain in a society that operates on the utility principle insofar as the principles of equal liberty and fair equality of opportunity are upheld. As he says, the utility principle "might sometimes lead, when consistently applied over time, to a basic structure securing the basic liberties; but there is no reason why it should do so in general."⁽³¹⁾ Thus, "citizens are to view themselves as free and equal persons;" and social institutions should be willingly complied with and recognized as just "in a well-ordered society governed by

29 Ibid.

30 Ibid.

31 Rawls, "Reply to Alexander & Musgrave". p.640.

either the principles of justice as fairness or the utility principle"⁽³²⁾. Now taking into consideration, certain social and economic inequalities that may exist in a society owing to such contingencies as family and class origins, natural endowments and the like, the question that must be asked is: what principle will appeal to free and equal persons who want the effects of chance to be regulated by some principle, if a reasonable principle exists? Rawls rules out desert as appropriate answer since "no one deserves his place in the distribution of talents, nor his starting place in society"⁽³³⁾.

In following the maximin criterion, "the distribution of abilities is viewed in some respects as a collective asset"⁽³⁴⁾. As mentioned earlier, even when an equal distribution of such assets would be more in keeping "with the equality of free moral persons, at least if the distribution were a matter of choice, this is not a reason for eliminating natural variations, much less for destroying unusual talents"⁽³⁵⁾. The unique rationality of the maximin approach lies precisely in its recognition of these natural variations. It allows institutions "to exploit the full range of abilities provided the

32 Rawls, "Some Reasons for the Maximin Criterion", p.144.

33 Ibid. p.145.

34 Ibid.

35 Ibid.

resulting inequalities are no greater than necessary to
produce corresponding advantages for the less fortunate".⁽³⁶⁾

It conforms to the precept cited by Marx "From each according
to his abilities, to each according to his needs".⁽³⁷⁾

How valid is Rawls' claim? Is the unique rationality of the maximin criterion so evident and so appealing such that reflective persons would generally want to adopt it as a principle to regulate their social fortune as well as their natural lottery? This does not seem to be the case, as several of Rawls' critics have effectively argued. For instance, Brian Barry argues that the principle may not even appeal to the rational contractors of the original position. If Barry is correct, we have at least a sufficient ground for rejecting Rawls' claim that the maximin rule is uniquely rational as a principle of just distribution. Let us examine the main features of Barry's argument.

The focus of Barry's criticism is the possibility of deriving the maximin rule from the conditions of the original position. The main issue here, he tells us, is "the rationality of committing oneself to a principle which ties the evaluation of a social system entirely to how well-off the 'worst-off representative man' is, however, he is de-

36 Ibid.

37 Rawls, "Reply to Alexander & Musgrave", p.654.

(38) fined". He asks whether it can ever be rational for the men of the original position "to espouse a principle which throws away so much prima facie relevant information, such as the average per capita income, the degree of dispersion round the mean, the distance between top and bottom and so on?" He recalls (39) an earlier unsuccessful attempt by Rawls to establish the maximin criterion in "Justice as Fairness". Here, Rawls argues that "the men in an informal approximation of the 'Original Position' would 'propose principles of a general kind' and that 'these principles will express the conditions in accordance with which each is the least unwilling to have his interest limited in the design of practices, given the competing interests of the others, on the supposition that interests of others will be limited likewise. The restriction which would so arise might be thought of as those a person would keep in mind if he were designing a practice in which his enemy were to assign him his place" (40). This argument, as Rawls himself subsequently acknowledges in A Theory of Justice, does not succeed because it amounts to reasoning from false premises. For the argument to succeed, as Barry points, it must be the case that one's enemy is indeed going to have the final

38 Brian Barry, The Liberal Theory of Justice, p.87.

39 Ibid.

40 Ibid. p.88.

word in assigning places in a common practice. Under such circumstances, it would seem reasonable for one to take adequate precaution to ensure that the worst-off position is as pleasant as possible. But the men in the original position do not expect their positions in life to be assigned to them by their enemies. Rather, they are aware that the allocation of places would depend largely on their personal characteristics and on luck. And since they do not possess adequate information on what personal characteristics they have, nor how their society will evolve and divide into classes, they can only anticipate a random process of allocation. Now, if the men are already aware of such a random process, why would Rawls expect them to believe and act as if their entire future is actually in the hands of their enemy? Is there no difference between getting into a situation by chance and getting into it through the machinations of an enemy?

The crucial point to note here is the problem of decision-making under conditions of uncertainty. Rawls' answer to it, as Barry points out, is the maximin criterion. It directs the decision maker to seek "the choice which maximizes the minimum pay-off"⁽⁴¹⁾. In effect, the decision maker must first look at the possible consequences of all the alternatives before him in order to discover the worse possible outcome that each might produce. He then picks the alternative

41 Ibid.

with the best possible consequence should everything go wrong. Obviously this is a conservative, play-safe strategy. But, as Barry asks, does it make sense as a universal response to uncertainty?⁽⁴²⁾

Barry's argument seems to carry considerable force. True enough, it is often said: "better a half a loaf of bread than nothing at all". But then, anyone who spends all his time and energy in search of "half loaves" when he could, or at least reasonably might get "full loaves", has tacitly condemned himself to the bottom of the social ladder for the rest of his life. It is quite in keeping with common sense that if a person already knows that he cannot ever make it to the top, then the wisest thing to do would be to settle for some place in between or try out other options. But to insist that one should always aim initially for the second best, as Rawls' conservative strategy appears to suggest, cannot be an adequate universal response to decision making under uncertainty conditions. It cannot be the most rational decision in the original position. If the men at the original position were to be sure that everyone else is out to damage the other person, then the strategy would be a rational response. But these men are innocently uncaring about what the other person is doing. Each one is only concerned to

42 Ibid.

mind his own business. They do not have the capacity of being envious. Hence some sort of strategy for playing the percentage can also be rational.

However, Rawls, while still retaining the same conclusions which he came to in "Justice as Fairness", has now abandoned the premises he used originally. He says explicitly in A Theory of Justice that the reference to one's enemy assigning one's place, does not establish the plausibility of his principle.

The persons in the original position do not, of course assume that their initial place in society is decided by a "malevolent opponent ... they should not reason from false premises". (43)

Rather, he argues that there are three features of any situation which increase the rationality of a conservative approach to decision making. He contends that "the original position manifests these features to the fullest possible degree, carrying them to the limit, so to speak". He lists these features as follows:

First, since the rule takes no account of the likelihood of the possible circumstances, there must be some reason for sharply discounting estimates of these probabilities.

Second, 'the person choosing has a conception of the good such that he cares very little if anything, for what he might gain above the minimum stipend that he can, in fact be sure of by following the maximin rule.

And third, the rejected alternatives have outcomes that one can hardly accept. (45)

43 Rawls, A Theory of Justice, p.153.

44 Ibid.

It turns out, therefore, that the unique rationality of the maximin approach may yet be established if these features could be shown to be well founded. But if Barry's objections are plausible it does not see that this could be done either.

It might be helpful if we restate the first feature with a concrete example. Let us suppose an agent, say Smith, is faced with a range of decisions X, Y and Z. While he does not know much about the probabilities attached to these decisions, he knows what the worse possible outcome of each of them might be. The rational thing for Smith to do in accordance with the maximin strategy, is to take the decision which he knows to have the most preferred worst outcome.

There is no doubt that under certain conditions of uncertainty this strategy might just be the one a reasonable person would follow. Consider a person who went to the race tracks with the only fifteen dollars he has in the world and wants to participate in the betting. Let us suppose he has several betting options. He could bet on a horse for five dollars, and should he win, he would win twenty-five dollars. He could bet for ten dollars with a possible win of one hundred dollars. And, he could throw in his entire fifteen dollars for a possible win of fifteen hundred dollars. Since

45 Ibid. p. 154.

the likelihood of his winning is entirely unknown, he would be unwise not to choose the worst preferred possible outcome. The wise thing for him to do if he really wants to bet is to bet with five dollars. Should he lose, he would still have ten dollars to hang on to.

The maximin strategy can thus be shown to be plausible under unique conditions such as this. However, Barry still argues that this sort of interpretation does not capture the crucial issue. He contends that the conditions of the original position are not the same as the ones presumed in the above example. The central issue raised by the "first feature" is whether it is rational under the circumstances of the original position "to be a 'pessimist' rather than an optimist, and to take account only of the possible worst outcome"⁽⁴⁶⁾. In other words, in any situation in which a person does not have adequate information it is always rational to adopt a playsafe conservative and indeed a pessimistic attitude.

Barry develops his objection along two lines. First he argues that there is no basis for rational choice in the original position in view of the constraints imposed by Rawls. If the hypothetical choosers do not know the probability of finding themselves in any of a number of different kinds of society, they would have to operate either on the principle

46 Ibid.

of "insufficient reason" -- "the idea that in the absence of information to the contrary, it is rational to assume that each separately - identifiable state is equally probable" -- or settle for a principle that does not require any information about probabilities.⁽⁴⁷⁾ The maximin rule would be a typical example of the latter. Barry argues that once we grant the possibility of a choice in which the parties have no information about probabilities, we must equally concede that they do not have any information on the "range of positions available, including of course, the worst-off position"⁽⁴⁸⁾. Hence, the choice at the original position is not one based on the principle of insufficient reason but a choice based on a principle that does not require information.

Although Barry does not see this argument as dealing a demolishing blow, it does raise a genuine problem with which Rawls should come to grips. If Rawls had taken the option of the principle of insufficient reason, it might be possible to rally up some defence for him. He rejects it, and, instead he chooses the maximin approach. Now, if the men are denied such vital information as that of probabilities, why should we even expect them to know of such particular positions as those of the worse-off. How does one simply know the worst-

47 Ibid.

48 Ibid. p.92.

off positions without knowing other probabilities? In the case of the betting man mentioned above, it is assumed at least that he knows what the possible outcomes can be.

Perhaps we need to take into account here what Rawls constantly refers to as the circumstances of justice: the conditions of moderate scarcity and the divergence of fundamental interests and ends, and a variety of opposing and incompatible beliefs. If the contractors are presumed to take these general facts into consideration, would this not enable them to know at least that some of them will inevitably end up with less than satisfactory positions in life? If everyone is going to struggle for as much of the common cake as he can possibly get, and the cake is not just enough for everyone to have his heart's content, then they could expect that some of them will necessarily come out worse-off than others. And, the maximin strategy would then appear as a reasonable means of preparing for such outcomes: a collective means of ensuring that the worse outcome should at least be tolerable.

But this consideration does not adequately answer Barry's objection. While the contractors may indeed know that some of them will end up in worse-off positions, this fact by itself does not suggest that they should not know the other probable outcomes. Recall the case of the man in the race track mentioned above. While he knows that his worse

outcome would be to lose his entire fifteen dollars and win nothing in return, he also knows what the best probable outcome would be for each decision he makes. Such probabilities help determine the rationality of his decision. For a person to ascertain whether a particular risk is worth taking or not, it is not sufficient for him just to know what the worse outcome will be. He needs to know the other probable outcomes including the best he can hope to derive from it. Also, in regard to the maximin approach as a collective means of ensuring that the worse possible outcomes are at least tolerable, this may indeed be a rational decision to make collectively. But it need not be rational as an individual decision. Consider once more the case of the man in the race track. It might be a valid prescription for one to say that people should not bet with all the money they have in their pocket when they are in the race track. As a ground for such a prescription we might cite the case of our man who has all the money he owns in the world in his pocket. But there may be others too in the same track, who can conveniently afford to bet with all the money they brought with them. For such individuals a conservative strategy can hardly be their most reasonable response to uncertainty. ⁽⁴⁹⁾

49 There is a strong tendency on the part of Rawls to use what might be collectively rational as a basis for asserting what is individually rational. It is not so evident that a direct implication relationship actually exists between the two. We will see more of this in the congruence argument.

The second feature listed by Rawls also raises some concern. Here, as Barry notes, Rawls assumes as a matter of psychological fact that "there is a definite threshold (and the same one for everybody) up to which increments of wealth and power are valued but above which they have little or no value."⁽⁵⁰⁾ One immediate problem implicit in this assumption is that of proceeding directly from a threshold for societal wealth to deduce the individual thresholds. Assuming, however, that this is only a practical difficulty, how does the assumption itself lead to the maximin rule? This is not so obvious. For instance, the choice of the contractors, as Rawls says, might be construed as a choice between different or alternative societies. But even then, in order for them to contemplate the choice of a society organized by the maximin rule, they need to know the minimum level of thresholds obtainable by individuals in that society and how this level coincides with the minimum level in which the threshold can effectively operate. Without such information, we cannot simply say whether any particular situation maximizing the minimum will be a good or bad idea. If the minimum obtainable threshold of primary goods in a society is below the level at which the threshold effectively operates, then it will be plainly irrational for anyone to desire such a society.

50 Barry, op. cit., p.97.

Barry's example of one hundred apples to be divided among ten people illustrates this point. If it so happens that these people have an identical Rawlsian threshold with respect to apples, then a strong case could be made for the maximin rule. Each apple up to the tenth would have considerable value for each person: and each apple after the tenth would have little or no value for each person. The maximin rule thus appears the best solution as it only requires that each person receives ten apples. Here emphasis is not on the total number of apples nor on the fact that giving ten apples to everyone reflects an equal distribution. What counts ultimately is the fact that each person receives his threshold for apples, which happens to be ten.

However, the case for the maximin rule becomes less convincing if we introduce another supposition. For instance, if we scale up the threshold for apples for each person to twelve, it is no longer obvious whether the best thing to do is to give each person ten apples. If we are mainly interested in equal distribution we may yet go ahead and give every person ten apples. But this hardly does any good to any of them because the threshold required for the fulfillment of their rational life plan is twelve apples. Now instead of employing the maximin rule and leaving everybody frustrated, why is it not rational to devise a method whereby we can guarantee eight of them twelve apples each? This appears to

be at least one rational approach; and, it is quite consistent with the principle of average utility. If it is tenable, it would seem to counter Rawls' claim that rational self-interested people will not choose the principle of average utility in the original position. And, that any principle that they choose must have roughly the same utility-function with respect to their wants. As Barry seems at least to have argued convincingly, it is not immediately obvious whether the maximin rule would always display the same utility function in every circumstance.

Rawls claims as third "special feature" of the original position "that the rejected alternatives have outcomes that one can hardly accept"⁽⁵¹⁾. The essence of this claim, Barry points out, is that "the outcomes of other choices may be extremely bad"⁽⁵²⁾. In effect, it would be irrational for the men in the original position to forgo the certainty of a satisfactory minimum for the chance of a worse outcome. Although Barry gives some rather interesting illustrations to support his objection to this claim, I shall dismiss the entire issue by recalling a more common place example in our local surrounding. Let us suppose there is a teller who works with the Canadian Imperial Bank of Commerce who is

51 Rawls, A Theory of Justice, p.154.

52 Barry. p.103.

thoroughly frustrated for lack of incentives on his job. On Rawls' account, it would be irrational for the teller to quit his job without first securing another job. Reason: it would be irrational for the teller to forgo the certainty of a satisfactory minimum (an unfulfilling job) for the chance of a worse outcome (unemployment). Well, if it is actually the case that the teller will not be able to find another job, (and he knows this to be the case), the best thing for him to do would be to stick to his bank. Here, the maximin rule is obviously justified. But unless the teller actually knows that he cannot get another job, why must he stick to his job. It is just as probable for him to find another job as it is for him to be jobless. Why is it so irrational to take a chance one way or the other?

Perhaps the discussion in this section could be brought to a close with one final remark. It appears that Barry has been quite successful in showing that the maximin rule does not uniquely follow from the original position. In other words, the circumstances of the original position are not such that we can say emphatically that if these men are rational they would always prefer the maximin rule to all other criteria. And, if these "representative" rational men can legitimately set aside the maximin rule, there is no reason why we should expect the criterion to appeal to actual

reflective persons in real life as being uniquely rational. The weakness or the inadequacy of Rawls' theory is all too transparent. The issue is not whether or not the maximin rule is reasonable. All through his argument, Barry does not fail to acknowledge those circumstances in which it would be rational to choose the maximin rule. Rather, the crux of the argument is whether there are situations in which it would be rational to set aside the maximin rule in favour of other sets of principles, for example, the principle of average utility. Rawls is committed to a defence of the unique superiority or rationality of the maximin rule over other sets of principles in the circumstances of justice and where moderate scarcity obtains. But this position, as Barry has successfully shown, is inconsistent with our ordinary considered judgments. Thus the maximin rule, like the principle of equal greatest liberty is essentially contestable. No adequate grounds have been given why rational human beings must be committed to either in the ways Rawls claims they must.

CONCLUSION

In this chapter we have examined the two principles of justice as fairness. These principles are quite familiar as Rawls says. The first principle requires that there should be an equality in the distribution of basic liberties. It forbids

the restriction of basic liberties except for the sake of greater basic liberties. The second principle with its maxim rule requires that if inequalities must be entertained in the distribution of social and economic goods, the inequalities must be to the advantage of the worst-off members of the community.

Although these principles may indeed be familiar as Rawls suggests, and many people, particularly in our society, do indeed consider them fair and reasonable, it is not obvious whether Rawls' lexical ordering of these principles would match all our considered judgments, even when the 'our' is rather circumscribed. A truly rational person may, by virtue of his rationality, waive his self-respect or autonomy in the interest of an economic or social good. Rawls has not been able to show why individuals must always rank their autonomy prior to every other good. Also, while the maximin rule may seem to favour those at the bottom of the ladder in the society, it is not equally just to the ones who could make it to the top. To justify the difference principle an appeal must be made to some utilitarian considerations. But Rawls would certainly have none of this.

It should seem obvious from this preliminary survey that every important aspect of Rawls' theory is essentially contestable. But is this not true of any philosophical work of major importance? Rawls' account is still one of very

considerable importance and magnitude. Even if he has not given us a workable theory, he has at least shown us a new way of looking at aged-long problems.

Anyone who is already familiar with Rawls' theory will doubtless agree that making a general survey of it is a risky venture that is bound to pose more questions than it answers. However, it should be pointed out that this survey does not make any pretence to thoroughness. It seeks merely to establish a general background for fairly limited purpose: to provide a framework for assessing Rawls' conception of the good. It does not propose to assess the overall argument of A Theory of Justice. Yet, having provided this background, the bearings of the various aspects of Rawls' theory of the good on the correctness of his theory of justice in general can be better appreciated. It should be pointed out too that even though the concept of the good may have been crucial to certain aspects of this survey, it has not been discussed in view of the full treatment it will receive in the subsequent chapters. Cross references will be made to these aspects in the course of the analysis of the good.

PART II

'GOODNESS AS RATIONALITY'

CHAPTER FOUR

THE PRIMARY GOOD

PREVIEW: THE ROLE OF THE GOOD IN JUSTICE AS FAIRNESS

Although Rawls is careful to lay out the major features of his theory of justice at the very beginning of A Theory of Justice, he does not discuss in any detail to what extent these features may be accepted or rejected independently of one another. He mentions rather casually that different "original positions" resulting in the derivation of different principles of justice could be conceived, even though his own original position remains, he contends, the philosophically most favoured interpretation of the initial situation that yields the principles of justice as fairness. Also, he concedes the possibility of one accepting his two principles of justice while rejecting the idea of their being derived from his own specification of the original position. Thus one may tend to view the various features in the theory as being

1 The possibility of arriving at different types of contract theories and each theory yielding a different conception of justice would seem to threaten the significance of the entire contract approach. Each contract theory might claim to have the most favoured interpretation of the initial situation, and hence the most plausible conception of justice. Rawls, however, does not anticipate any serious challenge from rival contract theories. He considers his original position as the most favoured interpretation of the initial situation because it leads to the principles of justice as fairness (principles which best accord with the considered judgments

independent, and capable of being investigated in isolation from one another.

There is no doubt that each of the main features in Rawls' theory is distinctly important. The truth of this claim becomes more apparent when one takes into account some of the major issues which currently dominate ethical theory. For example, the methodological process by which Rawls tries to establish the two principles of justice as fairness is an issue worthy of independent study regardless of what our particular intuitive reactions to the principles themselves are. His moral methodology may yet be acceptable either in part or as a whole even if we reject his two principles of justice. However, in view of the unique structure of Rawls' theory, one

or competent moral persons in reflective equilibrium). Also Rawls believes that his contract theory would be superior to other contract theories because the assumptions in his original position are a "constrained minimum". That is, while these assumptions are strong enough to yield a conception of justice, they are yet weak in the sense that they are widely acceptable. Neither of these two claims have been successfully demonstrated. If the objections raised in Chapter There are upheld, it does not appear that every rational person who is not already committed to the ideals of liberal democracy would endorse the two principles, at least not in the lexical ordering in which Rawls presents them. Assuming, however, that these objections are essentially mistaken, and that Rawls' principles could in fact be shown to be in accordance with the considered judgments of all moral persons, there is still the question of their justification at the original position. Why must we regard principles chosen by self-interested "egoists" as principles of justice? Rawls of course does not want his contractors to be referred to as egoists. But how else can we describe individuals who are said to be totally uncaring about others? Also, the claim that the conditions of the original position are a "constrained minimum" is not so evidently true. As we will come to see in this chapter,

runs a risk of misunderstanding the rationale of any particular feature in it, unless it is considered in the context of the entire theory. For example, even though we see the two principles of justice as fairness and the contract argument that led to them almost at the beginning of A Theory of Justice, we do not encounter the full justification that Rawls provides for these principles until we come to the arguments of reflective equilibrium and congruence. ⁽²⁾ Thus, it is necessary to consider the various parts of Rawls' theory as inter-dependent just as much for the understanding of the entire theory as for the understanding of the parts themselves.

This consideration would seem to justify the relevance of the survey of the essential framework of Rawls' A Theory of Justice that was conducted in part one of this study. To attain an adequate assessment of his theory of the good, we need first to understand some of the central themes of justice as fairness.

the assumption about primary goods may not be as minimally constrained as Rawls thinks. Again, assuming that Rawls can also be shown to be correct in the latter, that is, that the conditions of the original position are widely acceptable to reflective persons, might it not be the case that those very "controversial" issues that Rawls tries to avoid are the only ingredients that can yield a plausible moral conception? If morality is by nature problematic or controversial what is the special merit of a moral methodology that is unproblematic but incorrect?

² However, as Robert Paul Wolff rightly notes, it is still a difficult issue to decide the exact logical relationship between these three arguments. The three arguments do not cohere and the demands of each seems to undermine the requirements of the other two. See Paul Wolff, Understanding Rawls, p.191.

For, what Rawls does in his theory of the good is to relate the principles of justice as fairness to a theory of goodness as rationality. By appealing to what he calls "the principles of moral psychology" he tries to show how individual persons would, in ideal conditions, come to acquire the sense of justice. Rawls argues ultimately that a well-ordered society would permit individual autonomy, foster the realization of community, diminish the tendency to envy and selfishness, and enable human beings to express their nature as free and equal moral persons.

My approach in this part will follow the general outline set by Rawls in the development of the theory of the good. Emphasis will be on the following topics: (1) The theory of the primary goods; (2) the conception of rationality underlying the choice of primary goods; (3) the definition of a morally good person; (4) the desire to act justly; and (5) the good of the sense of justice. I proceed to these topics indirectly by way of a brief review of the role of the good within the general framework of justice as fairness.

One of the characteristic traits of A Theory of Justice is the attempt by Rawls to reverse or substantially alter the relationship between the concepts of right and good as these function in traditional ethical theory. Rawls maintains that the main lines of traditional moral philosophy have been dominated by theories in which the good is defined or as

a primitive conception (as in Moore) is taken to be prior to and independently of a conception of right; and the right is then defined as that which maximizes the good. Hence, those institutions and acts are said to be right "which of the available alternatives produce the most good, or at least as much good as any of the other institutions and acts open as real possibilities ..."⁽³⁾ Rawls contends that this approach easily elicits public support as it appears to manifest the common idea of rationality. It is natural, he says, "to think that rationality is maximixing something and that in morals it must be maximizing the good"⁽⁴⁾.

Rawls rejects the priority of the concept of good over that of right because of the consequences of this relationship. Apart from the fact that an independent concept of right does not emerge from it, it does seem to permit situations whereby individuals may be denied the freedom to implement a conception of good if it is incompatible with the dominant conception of good in a particular society.⁽⁵⁾ Should the good be allowed to set the stage for the definition of right, then what is right

3 Rawls' A Theory of Justice, p.24.

4 Ibid.

5 It is not so certain as we will come to see both in this chapter and in the next, whether Rawls himself is immune to this charge. He admits quite explicitly in "Goodness to Fairness" that the original position does favour certain conceptions of good over others.

might, under certain circumstances, implicitly define some putative good as wrong on quite arbitrary grounds. To avoid such situations, Rawls rejects the priority of good and sets out to provide an alternative theory that establishes an independent basis for the principles of right prior to the consideration of any substantive notion of good. Thus, he says that "the priority of the right over the good in justice as fairness turn out to be a central feature of the conception"⁽⁶⁾. This fact situates justice as fairness within a deontological theory of ethics in terms of conventional ethical theories.⁽⁷⁾

However, given this priority of the right, the good yet plays a significant role in justice as fairness. Rawls devotes as much as a third of A Theory of Justice to a theory of the good, showing how a minimal conception of the good is crucial to the derivation of the principles of justice and how the good is congruent with right and justice. In fact, on his account, it is a sufficient condition for rejecting any theory of right and principles of justice if these are

6 Rawls, op. cit., p.32.

7 This claim is challenged by Adina Schwartz. He argues that Rawls' assumption about the preference of more rather than less of the primary social goods is not morally neutral. It is based, he claims, on the preference of the majority of rational individuals. Hence justice as fairness is in fact a teleogocial theory of justice. This charge will be investigated in the course of this chapter.

incongruent with what he calls human good. Accordingly, in situating his account of justice within the deontological framework, Rawls is careful to note that deontological theories do not "characterize the rightness of institutions and acts independently from their consequences. All ethical doctrines worth our attention take consequences into account in judging rightness. One which did not would simply be irrational, crazy⁽⁸⁾".

Thus, Rawls assumes, in contrast with teleological theories, that "something is good only if it fits into ways of life consistent with the principles of right already on hand⁽⁹⁾". But to be able to formulate these principles however, he must still rely on some notion of good. Hence it is necessary, as he says, to make assumptions about the motives of the contractors in the original position. Even if these contractors are presumed to be rational, it remains the case that neither they nor any other human being for that matter, can decide on principles of justice or right without some prior knowledge of the good. But to keep these assumptions from jeopardizing the priority of the concept of right they must be restricted to "the bare essentials"⁽¹⁰⁾. Thus Rawls

8 Rawls, op. cit., p.30.

9 Ibid., p.396.

10 Ibid.

distinguishes between two theories of the good: the thin theory and the full theory. The thin theory is the minimal conception of the good used in arguing for the principles of justice. Its purpose is "to secure the premises about primary goods required to arrive at the principles of justice"⁽¹¹⁾.

Beyond such a minimal conception of the primary goods, the men of the original position do not have any other common information as to what is good or bad. They do not even share "a conception of the good by reference to which the fruition of their powers or even the satisfaction of their desires can be evaluated"⁽¹²⁾. Hence Rawls concludes that, if they are rational, they would not "risk their freedom by authorizing a standard of value to define what is to be maximized by a teleological principle of justice"⁽¹³⁾. Rather, they would agree "to an index of primary goods as a basis of interpersonal comparisons"⁽¹⁴⁾. The index still plays a subordinate role because the primary goods are just those things which "men generally want in order to achieve their ends whatever they are. Wanting these goods does not distinguish between one

11 Ibid.

12 Ibid, p.322.

13 Ibid, p.328.

14 Ibid.

(15)
person and another". And to accept them "for the purpose of
an index does not establish a standard of excellence". (16)

Thus, Rawls articulates the various roles of the thin theory in justice as fairness as follows:

First of all, it is used to define the least favoured members of society. The difference principle assumes that this can be done Further, the index of well-being and the expectations of representative man are specified in terms of primary goods. Rational individuals, whatever else they want, desires certain things as pre-requisites for carrying out their plans of life. Other things equal, they prefer a wider to a narrower liberty and opportunity, and a greater rather than a smaller share of wealth and income. That these things are good seems clear enough. But I have also said that self-respect and a sure confidence in the sense of one's own worth is perhaps the most important primary good. And this suggestion has been used in the argument for the two principles of justice.... Again, some view of goodness is used in defending justice as fairness against various objections. For example, it may be said that the persons in the original position know so little about their situation that a rational agreement upon principles of justice is impossible. Since they do not know what their aims are, they may find their plans utterly ruined by the principles to which they consent. Therefore, how can they reach a sensible decision?.... the thin theory of the good which the parties are assumed to accept shows that they should try to secure their liberty and self-respect and that, in order to advance their aims, whatever these are, they normally require more rather than less of the other primary goods. In entering into the original agreement, then, the parties suppose that their conceptions of the good have a certain structure, and this is sufficient to enable them to choose principles on a rational basis. (17)

15 Ibid.

16 Ibid.

17 Ibid., pp.396-7.

When the thin theory has been worked out and the index of primary goods accounted for, Rawls would now use the principles of justice deriving from it in the further development of what he calls the full theory of the good.

The first thing that we need to do here is to examine critically this thin theory of primary goods. Can it adequately fulfill all the roles listed above? Rawls tends to be rather casual about the thin theory. He believes that its main elements are "evident enough", and are held in common "by philosophers of markedly different persuasions"⁽¹⁸⁾. However, it is not at all certain whether the assumptions on which the theory is founded are so plainly true or justified or evident to all reflective people. In fact, there are good grounds, as Kai Nielsen argues, to believe that there are more difficulties in the theory than meets the eye and that Rawls' rather casual account of the primary goods may cause flaws in his theory that are not obviously correctable⁽¹⁹⁾. Therefore, in assessing the validity of the thin theory we will try to ascertain whether it is indeed the case that the claims it makes are so obviously true. Our discussion will focus mainly on two questions:

(1) Are the primary goods so morally neutral or value free such

18 Ibid, p.434

19 Kai Nielsen, "A CRITICAL NOTICE OF JOHN RAWLS' A THEORY OF JUSTICE". THE SECOND ORDER: AN AFRICAN JOURNAL OF PHILOSOPHY, 2 (2) 1973, p.105.

that all rational persons, regardless of race, colour or beliefs, must desire them if they are to be rational?; and (2). How adequate is the notion of rationality underlying the choice of these primary goods? The first question will be discussed for the remainder of this chapter, while the second question will be considered in the following chapter.

THE INDEX OF PRIMARY GOODS

Thus, we are to presume that there are certain types of goods (primary goods) that all rational men must desire, if they are to desire anything at all. And whatever their individual life plans may be, these primary goods are things that men generally want in order to achieve their life plans.

Regardless of what an individual's rational plans are in detail, it is assumed that there are various things which he would prefer more of rather than less. With more of these goods men can generally be assured of greater success in carrying out their intentions and in advancing their ends, whatever their ends may be. (20)

Rawls distinguishes between two basic types of primary goods: the primary social goods and the primary natural goods. "The primary social goods", he says, "to give them in broad categories, are rights and liberties, opportunities and powers, income and wealth"⁽²¹⁾. The primary natural goods, on the other

20 Rawls, A Theory of Justice, p.92.

21 Ibid.

hand, are such goods as health, vigour, intelligence and imagination. A very important good which Rawls does not discuss until much later in his work is the sense of one's own worth, that is one's good of self-respect. While this good is to be classed as a primary social good, Rawls, for simplicity, does not discuss it until Chapter VII.

Rawls does not say very much about the primary natural goods. Hence our discussion will focus accordingly on the primary social goods. ⁽²²⁾ Thus we might say that on Rawls' account, for a person to be rational and to want anything at all, he must want the primary social goods. And the contractors in the original position must want more of the primary social goods rather than less. They must show a preference for more of them rather than less. ⁽²³⁾ Perhaps one sure way of checking the truth of this claim is to examine carefully the various items he lists as primary goods, and see whether rational persons, no matter what else they want in life, would want more of these things if they are to be

22 Rawls' attitude towards the primary natural goods should not come as a surprise. To arrive at a unanimous choice of principles, Rawls must restrict the knowledge of natural talents as much as possible, as these could lead to biases of self-and group-interests. A choice made on the basis of natural endowment would be heteronomous and not autonomous.

23 The emphasis here is very important. Rawls' assumption is not merely that rational individuals prefer the primary goods, but that they prefer more of them, rather than less. This qualification is crucial to the question of justice. If some rational individuals prefer more and others prefer less, the

rational. We could begin the examination with the primary goods of wealth and income; these appear to be the least controversial items in Rawls' index.

Thus the question we need to ask at this point is rather straightforward. Is it the case that any reflective person, whatever his beliefs and plan of life might be, must want more of income and wealth if he is to be rational and want anything at all? Is it even the case that every reflective person desires income and wealth at all, not to say more or less of them? These questions cannot be answered in the affirmative if the terms 'income' and 'wealth' are understood in their ordinary connotations. It is true that concern for income and wealth is very pervasive in many Western societies which adhere to wage-labour system. But there are also other types of societies whose social organization does not require individual wealth and income. As Kai Nielsen points out, "income and indeed wealth, only go along with certain

question of just distribution may not arise as everyone may yet have enough to realize his life plan. But we must assume that rational individuals, under conditions of moderate scarcity, make equal demands for more rather than less of these primary goods. Even if it turns out in real life that some rational individuals will not desire more of these goods, owing to religious beliefs and the like, still it does not harm such individuals to desire more rather than less of them in the original position. An assumption of a preference for more rather than less of the primary goods is essential to the geometrical deduction of the principles of justice. One wonders of course whether this preference does not in fact sneak in some human ideal. This question will be closely examined later.

societies... they are not integral parts of the pattern of culture in every culture"⁽²⁴⁾. In Nigeria, for instance, there is a community of about twenty thousand people known as Aiyetoros. The social traits of this community bring to mind the picture of the utopia of Marxist Communism. The use of wage-labour for the purpose of organizing public labour is totally absent in this community. It is a classless community and the idea of acquiring individual wealth and income does not seem to cross the mind of its members. We might even bring the picture nearer to the North American scene. There are groups of individuals in this society who do not show any inclination towards personal wealth and income even though these items are very pervasive all through the society. There are the Gandhis, the romantics, the religious, the extreme pietists and the hyppies to mention but a few. Individuals in these groups are known to renounce wealth and income giving themselves up entirely to what they take to be higher levels of existence.

Now are we to understand that the Nigerian community mentioned above is irrationally organized or that its normal members are irrational? Also, can we say categorically that the hyppies, the Gandhis and the extreme pietists who do not desire wealth and income are all irrational? Perhaps it could

24 Kai Nielsen, op. cit., p.105.

be argued that socialist classless societies would be better off with wage-labour systems. Or, it may be shown that some members of the extreme religious sects as well as individuals in the hippie fraternities are irrational. But such an argument cannot be based on the fact that these individuals do not desire wealth and income. Nor can it be based on the mere fact that they belong to these unpopular organizations. Being religious can hardly be a stable criterion or even a reliable symptom for being irrational. Nor is it the case that all (25) people not concerned with income and wealth are religious.

Besides, Rawls cannot say that people should not be Gandhis, socialists or hippies. To defend such a proposition he would need a stronger conception of rationality than the one he employs at this stage of his theory. The notion of rationality in the thin theory does not stipulate any substantive ideals. An agent is rational provided he takes the most appropriate means (primary goods) toward the realization of his desired life plan. It is within such a framework that wealth and income must be seen as primary goods; goods that any rational person must desire in order to be rational and to want anything else. But how can this belief be justified if there are indeed certain individuals in various societies who do not desire these goods and who are not obviously irrational?

25 Ibid., p.106.

Are these goods, at best, anymore than the things generally sought in capitalist societies?

Rawls tries to deal with this problem in one of his subsequent replies "Fairness to Goodness". He notes that a number of critics, including Adina Schwartz and Michael Titelman, had attacked the original position on the grounds that it fails to treat persons fairly. In particular, they argue that his assumptions about primary goods tend to bias the choice of principles in an individualistic direction and against communitarian values. Thus they conclude that some people's conceptions of the good are unfairly favoured over others. In response to this objection, Rawls remarks that if the primary goods are correctly interpreted, the need to over-emphasize their individualistic bias does not arise. He argues that the "account of primary goods does not require that the index of well-being of social groups be computed solely in terms of what we think of as personal income and private wealth. Rather this index, he says, also "depends on the control over the stock and flow of means that people share in as members of associations and groups, including their role as citizens". Thus Rawls makes three distinctions which might justify income and wealth as primary goods. He says :

26 John Rawls, "Fairness to Goodness" Philosophical Review Vol.84 (1975), p.541.

27 Ibid.

- a) income and wealth is the legal command over the material means in general necessary to realize people's needs and interests, whether as individuals or members of associations, and the desire for such goods is not peculiar to a particular type of society (at any rate, not so long as the relevant class of societies is those with a sizeable population, fairly advanced technologically, and having a developed civilization and culture);
- b) income and wealth can be held in many forms public and associational as well as private and individual, (it is true that one develops the theory of the good from the notion of an individual's plan of life; but this does not imply that such plans must be individualistic; and
- c) the desire for income and wealth is distinct from the desire to be wealthy, and being wealthy is not a primary good. (28)

The interpretation of income and wealth in these passages may not be very familiar. Nonetheless, it should not be considered arbitrary. It is a standard usage in economics, and Mill's definition of income and wealth is along this line. Assuming its validity therefore, we shall try to see whether the charge that Rawls' conception of the primary goods is ethnocentric or dependent on the particular values and mores of capitalist societies can be effectively dismissed.

28 Ibid, p.542.

In regard to the distinction between "wealth" and "wealthy" Rawls maintains that there is a relational sense, not of the term "wealth" but of the adjective "wealthy". He explains that someone who is wealthy, or who aspires to be so, is someone who has, or strives to possess, far more wealth than the average person. Relative to others, his command over the means necessary to attain human ends is very great.

It would seem from Rawls' clarification that income and wealth need not be seen as something unique to some particular forms of societies or to certain kinds of values. While they may still be considered as some of the basic features in a capitalist society, they may not be in conflict with the classless structure of socialist-oriented societies. If wealth consists in the legal command over such items as food and land, buildings and machines, and rights to use or to receive or in any way to derive benefit from such items, it is difficult to see how any individual or groups can survive without wealth. And, unless individuals are presumed to desire wealth in this sense at least, the planning of any economy whether capitalist or otherwise cannot even begin. The needs and interests of individuals together with the means of attaining them must first be ascertained in order to plan an economy. Thus, as Rawls notes:

both Marx and Mill recognize a basic kind of motivation that persists through the course of historical change. The desire for income and wealth (income being a flow, wealth a stock), understood as the (legal) command over exchangeable means for satisfying human needs and interests. (29)

Although the general idea of Rawls' interpretation seems quite tenable, his definition of "wealth" still needs to be made more precise. As it stands, the definition does

29 Ibid, p.541.

not bring out an essential distinction between what economists generally refer to as "productive" and "non-productive" wealth. This distinction is important, if wealth must be considered as a primary good. For instance, if Rawls' definition is taken in the sense of productive wealth -- wealth which is sought for the purpose of securing greater income -- its validity as a primary good for all rational persons remains essentially contestable once more. As was argued before, in an ideal socialist state where there is no private ownership of the means of production it would be irrational for anyone to seek wealth as means of procuring greater income. Also, the notion of wealth in a productive sense cannot be a primary good for the hippie, the Gandhi or even the Roman Catholic priest who takes the solemn vow of poverty. While each of these characters may indeed need all the items listed by Rawls -- food and land, buildings and machines -- they do not need them as means for securing income. Rather they need them as "non-productive" wealth, that is wealth sought for the purpose of satisfying human needs and interests. In the latter sense, whatever may be properly defined as capable of satisfying human needs may be called a primary good. Thus, wealth and income may be described as primary goods -- goods to be desired by all rational persons -- only if wealth is understood as defined by Rawls but in a non-productive sense.

Now assuming that wealth and income may thus be regarded as primary goods, the question yet remains: why must we also assume that rational individuals share a preference for a greater rather than a lesser amount of wealth? This is the crux of the criticism raised by Adina Schwartz. As mentioned earlier, Rawls wants to make the assumption that rational individuals have a preference for a greater rather than a lesser amount of primary goods so that he can proceed deductively to the principles of justice. But is such a preference so obviously true? As Schwartz points out, a socialist might object to Rawls' claim. The latter could claim that "his good is furthered by just enough wealth so that he is decently fed, housed, and clothed"⁽³⁰⁾. Of course, Rawls might respond to this objection by drawing attention to the veil of ignorance. He could argue that since it is

³⁰ Adina Schwartz, "Moral Neutrality & Primary Goods", Ethics 83, 1973, p.304. Again one might respond here that Rawls merely intends the constraint of desiring more rather than less of the primary goods, as a strategy for securing the principles of justice. Hence in a well-ordered society some rational individuals may not desire more of these goods. Rawls in fact argues along these lines in "Fairness to Goodness". However this argument does not help Rawls. His attitude to the assumptions in the original position is much stronger than what this argument suggests. Ordinarily one would think that since the original position is merely an hypothetical device, the assumption therein would be arbitrarily drawn. But Rawls does not see it that way. He argues that the constraints in the original position are peculiar to moral theory. Besides, the purpose of the thin theory is to show that the assumptions about primary goods have a basis in human psychology (they are "generally facts" about human nature). Now if this is indeed Rawls' attitude, it is irrelevant to make a distinction

practically impossible for anyone to know either the specific amount of primary goods that would further his good, the most prudent thing to do is to desire a greater rather than a lesser amount of the primary goods at the original position. Besides, if it turns out that the good of the socialist can actually be realized by a lesser amount of wealth, he need not accept a greater amount of wealth. Also, Rawls might argue that the socialist would not be harmed by a distributive structure which is based on a preference for a greater amount of wealth; he could always refuse to accept the wealth he does not need or even donate it to charities.

But the socialist may yet be able to counter Rawls' argument. He could insist, as Schwartz points out, that living in a society based on a preference for a greater rather than a lesser amount of wealth is harmful to him. He might tend to devote much of his time to material pursuits and his philosophic or artistic aspirations would be in jeopardy. He might not be able to form strong ties of affection with other human beings since most people in such a society would be more interested in wealth rather than in people. And, as a result of competition for more wealth and income envy and

between the standpoint of the original position as against that of a well-ordered society. The motivational assumptions about the rational contractors must also be true of reflective persons in a well-ordered society. This point will be examined further in subsequent pages.

inter-personal rivalries would be prevalent in the society. By contrast, the socialist could argue that a distributive structure based on a preference for a minimal amount of wealth would encourage close personal ties as well as a more artistic way of existence. Thus he could claim that the assumption of a preference for a greater amount of wealth is against his interest and that a system based on a preference for less wealth is the one that furthers his life plan and his conception of the good.⁽³¹⁾

It is not so evident how the interpretation of income and wealth from Rawls' "Fairness to Goodness" can answer these arguments. What bothers the socialist is not the claim that he needs these goods, but that he needs more rather than less of them to be rational. Is it not rather the case that he will be irrational to desire more income and wealth when he knows that such a desire is incompatible with his life plan? Perhaps Rawls might want to circumvent this difficulty by claiming that in our society most persons have a preference for more income and wealth. And, even if it is true that some rational individuals do not have a preference for more of these things, his theory is still successful in as much as it captures the preference of the majority. The assumption about the primary goods is made from "general facts" of psychology. What most people want most of the time, and not what everyone wants all the time. The following passages in the final

31 Ibid.

chapter of A Theory of Justice tend to support this line of reasoning.

In one of his final remarks on the good of the sense of justice, Rawls admits that in a well-ordered society there may be "some persons for whom the affirmation of their sense of justice is not a good. Given their aims and wants and the peculiarities of their nature, the thin account of the good does not define reasons sufficient for them to maintain this regulative sentiment"⁽³²⁾. But Rawls insists that such individuals do not pose a threat to his theory. He argues that "to justify a conception of justice we do not have to contend that everyone, whatever his capacities and desires, has a sufficient reason (as defined by the thin theory) to preserve his sense of justice. For our good depends upon the sorts of persons we are, the kinds of wants and aspirations we have and are capable of"⁽³³⁾. While these claims seem reasonable and quite consistent with the fallibilist approach, they force Rawls into a rather awkward position. If there are persons who do not have sufficient reasons (as defined by the thin theory) to affirm their sense of justice, then we could also say that there are rational persons who do not have a preference for a greater amount of the primary goods. (It should

32 Rawls, A Theory of Justice, p.575.

33 Ibid, p.542.

be pointed out of course that the persons Rawls is referring to primarily in these passages are the amoralists.) Yet if the latter is true, and Rawls expects each member of a well-ordered society to affirm the public conception of justice, would the life plans of some individuals not be in jeopardy? Evidently, the majority is well protected because the principle of justice to be publicly affirmed took into account their preference for more rather than less of the primary goods. But how do we account for those rational individuals who do not share this preference? Individuals such as the socialists mentioned above? Of course, Rawls cannot say that a person is irrational if his life plan does not involve a preference for more rather than less of the primary goods. To say this would involve him in an appeal to some ideal of rationality. (The notion of rationality in the thin theory must be morally neutral.) Does this not mean then that, the derivation of Rawls' two principles of justice is based on the assumption that most rational individuals share a preference for more of the primary goods, as distinct from the assumption that all rational persons share a preference for them? And if this is the case, is Rawls not committed to claiming, as Schwartz argues, that "the majority's conceptions of the good should determine the choice of the principles of justice?" We will

34 Schwartz, op. cit., p.306.

return to this question once more after reviewing other items in Rawls' index of primary goods.

Other items in Rawls' index of Primary Goods include rights and liberties, powers and opportunities, and the good of self-respect. At first sight, these items appear to be sufficiently neutral and so broadly based that any rational man anywhere and anywhen cannot but desire them if he were to desire anything at all. But as soon as we try to give them specific content, our optimism begins to dwindle. For instance, what sort of rights, powers, liberties and opportunities are we to define as primary goods? Here we need to be mindful of the fact that in political theory lists of natural rights and declarations of natural rights have varied from place to place and over time. ⁽³⁵⁾ Within the traditions of liberalism itself, there are some individuals who emphasize political liberties while others insist on liberty of conscience. In face of such a diversity therefore, is it ever possible to translate these concepts with some specificity such that a particular set of objects X, Y, and Z can be described as primary goods that all rational persons desire?

Perhaps this problem could be mitigated, if, as Nielsen suggests, we define these goods as basic liberties such as "the right to vote, to stand for public office, freedom of speech

35 Kai Nielsen, op. cit., p.107.

and assembly, liberty of conscience, freedom of thought, freedom to hold personal property, i.e. to have what Marxists call consumer durables"⁽³⁶⁾. While such a definition does indeed provide some content, we are immediately faced with the problem of determining whether in actual fact these are the things that all rational men must want, regardless of their particular situations. Without even questioning the scope or the exhaustiveness of the list, it seems quite obvious that not all rational men cherish the right to vote. In the United States of America, as it is in all the so-called democratic countries, only a fraction of the population goes to the poll to elect new leaders. In fact, the campaign directors for the democratic candidate in the 1976 U.S. elections were gravely concerned about the general apathy of the Americans towards the poll. Now in face of such a glaring fact, how can one maintain that the right to vote is a primary good that all rational men must desire. Of course, it could be argued -- and not implausibly -- that the fact that people do not generally go to vote does not mean that they do not cherish the right to vote: "wanting" the right to vote is not the same as "exercising" the right to vote. We may yet prove our point from another example while conceding this distinction. According to Nielsen, Enzenberger in his Study Das Verhor Von Habana "makes it

36 Ibid.

perfectly evident that there are many rational present day Cubans who do not at all want the right to vote: they do not want electoral politics or representative democracy".⁽³⁷⁾ One might add here too that even such a thing so basic as freedom of thought has not been wanted by all rational men in all domains and at all times. There were strong arguments against it during the Middle Ages. And even today, when one thinks of the magic appeal of the image of the Grand Inquisitor in The Brothers Karamazov, there is no reason to believe that all rational persons would be unanimous in desiring freedom of thought.

The picture does not look any better when we come to examine self-respect. While it is true that a lot of people in Western societies do place a high premium on the notion of self-respect, it is far from being obvious whether all rational persons see it as a supreme primary good that must be sought at the cost of everything else. It is not uncommon in our society to find individuals who are totally uncaring about others and who have no traceable impulse of self-respect. Perhaps such individuals could be shown to be irrational in some 'deep sense' of the term "rational". But what are we to say about a thorough amoralist or what Nielsen calls "a

37 Ibid.

calculating, rational egoist" who can only see any concern for self-respect as a "moralistic hang up". Indeed such an egoist could be made to see how mistaken he is in his viewpoint. But if we must describe him as being irrational, we would need a stronger conception of rationality than the one on which Rawls bases his thin theory. As was mentioned above, Rawls wants to use "rational" in a morally neutral manner; he wants to use the term the way economists use it, that is taking the most effective means towards a desired end. With such a rather thin conception of rationality, it is difficult to see how anything at all could be strongly defined as a primary good which all rational persons must want, without first assuming one overriding end that all men must want if they are rational. Apart from the case of the amoralist, there are Christians and other religious believers who seek God as the greatest and ultimate end of human existence. Many of these, following the example of Christ, have freely detached themselves from all worldly concern. They have, as the Bible put it, emptied themselves and taken on the form of "slaves" in order to reach God. To them what Rawls eloquently speaks of as the good of self-respect is just one variant of the sin of pride which stands in the way of accepting the reality of God. Now, even if we are inclined to regard these people as irrational, this cannot be done on Rawls' theory of rationality.

It seems evident from these remarks that there are at least two important criticisms against Rawls' index of primary goods: the lack of specificity and the inadequacy of the theory of rationality underlying the choice of these goods. The central assumption of the thin theory is that there are certain goods -- primary goods -- of which all rational men would desire more rather than less if they desire anything at all. But if the considerations raised above are valid, it does not seem to be the case that there are many specific things that we can quite uncontroversially claim that any rational man must want if he wants anything at all. To make such a claim, we should at least recognize, as Nielsen indicates, that the "must" in our proposition has the illocutionary force of a moral commitment. But this in turn would be incompatible with Rawls' deontological commitment and fit in badly with his contractarianism. The primary goods as well as the principles of justice as fairness must be able to generate their own support.

However, in spite of all the difficulties raised so far, the assumption about the preference for the primary goods from what Rawls refers to as "general facts" of psychology may yet seem to have a strong appeal. If we look around us in our society here, we could easily come to agree with Rawls, that it is indeed the case that most persons with rational plans of life do in fact have a preference for these primary

goods. Such signs as "BINGO ON FRIDAY NITES" in front of parochial churches is a constant reminder that some material income and wealth is indispensable even in such sacred domain as the church. However, Rawls' assumption runs aground in view of what he wants to get from it. To be able to derive his principles of justice from the thin theory, he would need to make the assumption that all rational persons have a preference for all these goods. But as indicated earlier, it might be the case that Rawls in fact only associates this preference with most and not all rational persons. Why would he not still derive his principles from the assumption, if it is in fact true only of most persons. There are two straightforward answers to this question: (1) If Rawls bases the derivation of his principles on the preference of most rational individuals, then, like those theories he describes as "teleological theories", he himself would be saying that the majority's conceptions of the good should determine the principles of justice. Rawls would thus be hard pressed to explain why it is more desirable to choose the principles of justice on the basis of a preference common to the majority and not on the basis of a peculiarly virtuous person's conception of the good. (2) Such a move cannot yield the Archimedean point to judge the whole of morality. Rawls would not be doing much

38 Schwartz, op. cit., p.302.

more than simply telling us that when the majority's conception of the good is such and such, then such and such principle would apply. This form of argument falls short of the type of justification the sceptic is after.

Perhaps Rawls might yet want to make a last ditch stand. He could argue that most people have a preference for his primary goods and the few individuals who do not, are just the odd ones. Since the majority of the members of society cannot be like these individuals, why could he not use the preference of the majority as a starting point? Such a move would involve Rawls in a stronger conception of rationality than he wants to use in the thin theory. Above all, he must show that it is only the majority's conception of the good that truly reflects the general facts of human psychology. Perhaps this could be done. One might be able to show that the minority's conception of the good deviates from the normal order of human nature. Arguments along this line have been advanced against such practices as Homosexuality, Apartheid etc. However, this approach inevitably involves Rawls in stipulating ideals of rationality and moral personality and simply accepting the majority's view as determining what is right. And should he take to this, the claim that the conditions of the original position are a constrained minimum would be insupportable.

From the discussion in this chapter therefore, it would seem that one of the central claims in Rawls' theory of the good -- the claim that rational persons must have a preference for his index of primary goods if they are to be rational -- is not obviously true. There are persons who might be rational even if they do not have any preference for these goods. Thus, the effectiveness of the theory of the primary goods as one of the central premises in the derivation of justice as fairness is at best contestable. We shall now turn to the theory of rationality underlying the choice of these primary goods.

CHAPTER FIVE

"RATIONALITY" IN THE THIN THEORY OF THE GOOD"

Rawls defines the primary goods as those things which any rational person would want to have in order to fulfill his life plan. Thus we might presume that each of the contractors in the original position would want to have at least some of the primary goods since they are said to be rational. But at least one question remains: how are these contractors to know the relevant primary goods that would satisfy their life plans when they do not know what these plans are? (Recall that these men are placed behind the veil of ignorance, hence none of them knows the particular things he wants in life.) The thin theory of the good is Rawls' answer to this question. In the thin theory, Rawls tries to show that even though different people might want different things, there are certain standard features of wants which provide the basis for arguing to standardized means of want-fulfillment.

Rawls makes two important moves in the thin theory. First he attributes to each person a "rational plan" which defines his main expectation in life. He then goes on to define a person's "good" as the successful execution of a rational plan of life. Thus the first aim of the thin theory, as he says, "is to provide a criterion for the good of the

person. This criterion is defined chiefly by reference to the rational plan chosen with full deliberative rationality⁽¹⁾".

Although the thin theory contains several thought-provoking features, the discussion in this chapter will concentrate on the conception of rationality that emerges from it. As Brian Barry rightly observes, the notion of rationality turns out to be the centre-piece of the notion of primary goods as well as the thin theory of the good on which it rests. For instance, if we eliminate the word "rational" leaving only the word "plan", and if we also ignore the "planning" and "calculating" overtones of the word "plan", we are left essentially with the claim that a man's good is to be conceived as consisting in his getting more rather than less of whatever things it is that he happens to want. In which case, if a man happens to want cocaine, then having more rather than less of cocaine must be conceived as his good. Obviously such a crude proposition cannot correctly be attributed to Rawls. Thus much of our attention will have to focus on the concept of rationality. Whatever we are able to extract from Rawls' understanding of this concept will be extremely important not only in the assessment of the thin theory but the entire thought-experiment of the original position. Rawls wants to

1 Rawls, A Theory of Justice, p.423.

2 Brian Barry, op. cit., p.27.

maintain that there is a connection between rationality and the maximization of benefits. The representative persons in the original position are to be considered rational if their choice of principles would yield such outcomes in virtue of which each person would receive a full share of the social benefits. In other words, no person or groups should be better off at the expense of others. Each representative person must be satisfied that the outcome of their collective choice is the best he could receive in society without causing any other person to receive less benefits. Now to be able to adjudicate effectively on this claim we need to scrutinize Rawls' conception of rationality. We need to be satisfied that the principles of rational choice which he attributes to these persons are indeed correct.

Another important consideration why Rawls' conception of rationality calls for a careful scrutiny is his overall program of attempting to reconcile morality with rationality. As David Gauthier rightly notes, one of the central issues in modern moral philosophy is the reconciliation of morality with rationality. Those philosophers who believe in the possibility of normative ethics as a viable and objective discipline invariably insist that rationality is an essential element in morality. But when it comes to the task of the actual reconciliation of the two concepts, the prospects of this position no longer look attractive. Some philosophers advance a formula

selecting a conception of rationality and a conception of morality so that the two would fit. But what is of greater interest and value, according to Gauthier, is to attempt the reconciliation beginning with our intuitive conceptions of rationality and morality. It is the latter that Rawls claims to have accomplished in A Theory of Justice.⁽³⁾ He tries to show how the conception of rationality commonly held in social theory (a conception that does not appeal to any ideals of human excellence or moral personality) could lead a person to affirm his sense of justice. It remains to be seen in the following chapters whether Rawls has actually succeeded in doing this, or even whether anyone could ever succeed in doing it. Thus, the discussion in this chapter will highlight such aspects in the thin theory that are relevant to the notion of rationality. An appropriate starting point might be Rawls' three stage definition of goodness.

(4)

THE THREE STAGE DEFINITION OF GOODNESS

Rawls does not proceed directly to apply the concept of rationality to the assessment of plans. Rather, he first

3 David Gauthier, "Justice and Natural Endowment: Towards a critique of Rawls' Ideological Framework", Social Theory and Practice, (3), (1974), p. 3.

4 It is important to note that when Rawls talks of providing a definition for goodness, he should not be understood as intending to provide a definition by way of concept

illustrates his conception of it by considering what he terms "Simpler Cases". He says:

Thus I suppose the definition to have three stages as follows (for simplicity these stages are formulated using the concept of goodness rather than that of better than): (1) A is a good X if and only if A has the properties (to a higher degree than the average or standard X) which is rational to want in an X, giving what X's are used for, or expected to do, and the like (whichever rider is appropriate); (2) A is a good X for K (where K is some person) if and only if A has the properties which it is rational for K to want in an X, or whatever; (3) the same as (2) but adding a clause to the effect that K's plan of life, or that part of it relevant in the present instance, is itself rational. (5)

Although what the notion of rationality means when it is applied to plans is not conveyed in this definition -- Rawls promises to deal with this aspect later on -- we are to understand that once an object is known to have "the properties that it is rational for someone with a rational plan of life to want, then we have shown that it is good for him. And if certain sorts of things satisfy this condition for persons generally, then these things are human goods"⁽⁶⁾.

The notion of rationality that emerges from this definition identifies rationality with utility-maximization:

identity. Rather, what he refers to as definition is no more than "a rough guideline for constructing substitute expressions that can be used to say what on reflection we want to say." (Rawls, p.401).

5 Ibid, p.399.

6 Ibid.

an individual is expected to have a genuine desire to maximize his expectable utilities if he is to be considered rational. One important question comes to mind immediately with regard to this definition. How does such a morally neutral conception of rationality relate to the concept of morality? For instance, how does a person's prudential concerns relate to his goodness as a moral agent? Although this question is quite significant, it is rather premature to begin to discuss it at this point without first seeing other aspects of Rawls' conception of rationality. Nevertheless, it does provide a good insight on the nature of the task that lies ahead; namely, the feasibility of using a utilitarian maximizing conception of rationality as one of the essential premises for generating a non-utilitarian theory of justice or morality. This question will be discussed in some detail later on in this chapter. Meanwhile we will examine some of the specific claims of the three stage definition.

The three stage definition of goodness makes two important claims: first it assumes that an account of rationality could be given in such a way that its objective characteristics may be plainly recognized. This would mean that in any given situation, it is possible to specify a range of objects that would be rational for rational people to want regardless of such contingencies as individual feelings, beliefs and attitudes. Insofar as we are able to establish

that an object has the properties that it is rational for someone with a rational plan of life to want, then we have shown that object to be a good for him.⁽⁷⁾ An agent can only be seen to be rational under such circumstances if he desires the object.

There is no doubt that certain aspects of this definition can be shown to be well-founded. Indeed, Rawls is correct to note that "there is wide agreement among philosophers of different persuasions on an account of the good along these lines".⁽⁸⁾ The first part of the definition portrays the sense in which an object is said to be good when it displays the features that are commonly desired in objects of its kind. One could claim rather objectively and uncontroversially that any given object is a good one, so long as there is enough similarity of interests and circumstances among the persons concerned with objects of its kind. Consider such artifacts as snow tires, cricket bats, winter jackets,

7 The term "rational" as it applies to life plans here is to be understood in terms of the achievement of a person's wants considered within the framework of the overall system of his goals and desires. It does not stipulate any moral ideal or any ideal of rationality. The only relevant consideration is whether a person's wants are consistent with the overall system of his desires and goals in life. How adequate is this notion of rationality? This is one of the main issues to be determined in the course of this chapter.

8 Rawls, op. cit., p.400.

hockey sticks and the like. Given the constraint in the first stage of Rawls' definition, we can say rather objectively that a particular object in this range of objects is a good one. In the case of hockey sticks, for instance, such properties as size, weight, length, durability and the like would be considered relevant by hockey players. Hence for any hockey stick to be considered good, it must have these properties to the extent that they are desired by hockey players.

Perhaps it should be pointed out here that even though Rawls' definition may be properly described as an objective characterization of the good, it is nonetheless contextual. Rawls says quite explicitly that whenever we make the judgment that a particular object is good, there "stands in the background a point of view from which ... (the object) is being appraised, although of course this point of view need not be made explicit"⁽⁹⁾. We can always ascertain and characterize this point of view by "identifying the persons whose concerns are relevant for making judgments, and then,⁽¹⁰⁾ by describing the interests which they take in the object". Thus, a good cricket player we might say is one who has the skills and abilities that we naturally expect to find in

9 Ibid, p.402.

10 Ibid.

cricket players. A bad cricket player, by contrast, would be one who lacks such skills and abilities. Rawls also adds that "the point of view varies from case to case and the definition of goodness contains no general formula for determining it. These matters are to be explained as the occasion arises or gathered from the context".⁽¹¹⁾ Should we take this last remark seriously, it becomes necessary to be cautious about what we attribute to Rawls. While his account of "goodness as rationality" seeks to present the good as an objective concept, he is not involved in a program of the Platonic sort. He does not claim that if a thing is good, it has to be good in every conceivable world whether rational persons exist there or not. Rather, for a thing to be good, it must have the natural properties that rational persons normally expect it to have.⁽¹²⁾ And since rational people are products of specific environments their expectations vary from place to place. Whatever be the case, the goodness of an object is always contingent upon the point of view of particular rational agents.

11 Ibid, p.403.

12 There is of course the question of how we are to determine which of the persons in any given situation is a rational person.

The second part of Rawls' definition does not pose any particular problem. It captures the sense in which an object is said to be good for a person since it achieves what he desires in objects of its kind. An example here would be a person who desires cars of high-speed performance. Any car that has high speed performance can be said to be a good car for that person.

While the first two stages of Rawls' definition thus seems uncontroversial, their scope is rather limited. As Rawls himself points out, "no questions are raised about the ratio-
(13)
nality of the ends". Since it is normal to assess the rationality of an individual's desires, it would be necessary to extend the definition beyond these two stages. Rawls sees the need and the third stage of his definition tries to cater for this aspect. Thus in the third stage of his definition the question of the rationality of an individual's desires emerges. This part of the definition articulates that sense in which an object is said to serve a person's good when it achieves what he wants in objects of its kind, and his desire in the object is part of a system of desires and aims which correctly reflects what he wants in life. Here we might think of a race driver whose life plan is to win the "INDY 500". Whatever enables him to win this race may be referred to as his good.

13 Ibid, p.407.

Thus, before we come to say that a particular plan is a good plan of life, we must first ascertain that it is a rational plan. It is the rational plan for a person that ultimately determines the good for him. This viewpoint may be formulated in a proposition as follows: for any given life plan, of any given person, if the plan is a rational one, then the person's conception of his own good is a rational one. And, if a person's conception of his own good is a rational one, then his dominant interests, aims and desires also become rational. In such circumstances, one's apparent good may be said to coincide with his real good.

Evidently, the first two senses of good in Rawls' definition are instrumental. As Adina Schwartz notes, when a person appeals to the notion of good in evaluating an object in any of these two senses, the relevant question to ask is whether the object achieves its designated purpose. (14) The value of the purpose itself is hardly questioned. One is not committed to any particular criticism of ends when appealing to either of these two senses. Hence we might say that the first two parts of Rawls' "rationalistic" definition of good is indeed morally neutral. Rawls seems to be able to appeal to a purely instrumental sense of rationality in these two parts.

14 Adina Schwartz, op. cit., p.299.

However the claim of moral neutrality does not seem so easily defensible in the third part of Rawls' definition; at least such a claim does not seem to follow from the assumptions in the first two parts of his definition. This part allows a situation whereby an object may indeed have all the properties that a person might desire in it in view of his desires in objects of its kind. But to assert that the object serves his good (furthers his rational life plan) his desires in objects of its kind must be part of rational plan of life. An illustration might be helpful here. A drug addict might desire a particular type of drug, say cocaine, because cocaine achieves the various properties he desires in drugs. But since his desire in drugs is obviously not a part of a rational life plan, cocaine cannot be defined as his good even if the latter does have all the properties he desires in drugs. Now, while the issue involved here seems rather simple and obvious, one is not so sure at first whether Rawls can consistently defend a claim such as this without at least an implicit normative appeal. If the desires a person has in an object must be coherent with a rational plan of life for the object to be his good, and a plan of life is not rational just because a person happens to desire it, does it not mean then that there is some ideal beyond the factor of want that determines the "good" of a person? If so, does the third part of Rawls' definition not entail a normative appeal?

The third part of Rawls' definition apparently entails a criticism of ends. But, as Schwartz argues, this fact does not constrain Rawls to stipulate a moral ideal. He suggests that Rawls might yet be able to use his definition of the first two senses of "good" to establish the moral neutrality of his definition of the third sense of "good". Rawls' central assumption in the latter follows from:

Royce's thought that a person may be regarded as human life lived according to a plan. Therefore, according to the second part of the definition, a person's good has the properties which it is rational for him to want in a plan. Since part three states that an X is a good X if and only if it satisfies part two of the definition and fits into a rational plan of life, Rawls obviously thinks that rationality is the property that it is rational to want in a plan. (15)

Schwartz supports this claim by considering the purpose of a plan:

It seems evident that a plan is an ordering of desires and goals for the purpose of getting these achieved. Therefore, it is rational (in an instrumental sense of rationality) to want a plan to have the property which leads to the best achievement of desires and goals. In saying that a man's good is a rational plan of life and that an object is good for a person only if it "has the properties that is rational for someone with a rational plan of life to want", Rawls is not stipulating a moral ideal. He is simply clarifying the commonly held notion that a person's good (at least in a nonmoral sense) is the achievement of his wants considered within the framework of the overall system of his desires. (16)

15 Adina Schwartz, op. cit., p.300.

16 Ibid.

This passage seems to present an accurate reading of Rawls. Perhaps if we recall the fact that Rawls' thought has been considerably influenced by the rationalistic philosophy of the enlightenment, we would easily come to see the force of Schwartz' claim. However, the assumption that rationality is itself the property that it is rational to want in a plan is very contestable. If, as Schwartz says, the purpose of a plan is the ordering of desires and goals so that these may be achieved, how does rationality (which is the ordering exercise) again become the property that it is rational to want in a plan? Does the term "plan" not always suggest planning "for" something? It is indeed correct as Schwartz points out that a person's good is sometimes defined in terms of the achievement of his wants considered within the framework of the overall system of his desires. But if this definition is morally neutral what prevents us from defining a man's successful attainment of cocaine as his good, if the overall system of his desires happens to centre on drugs, and cocaine achieving all he wants in drugs? It would seem also that if the fact of a match between a particular want and the overall system of a person's desires is sufficient to define his good, then it would be pointless to ask whether a plan that actually attains this match is yet the good for a person. But we can always say without contradiction that even though a particular attainment reflects what a

person really wants in life, yet it is not good for him or for anyone else to want it. Consider the case of someone whose overall system of desires is geared toward receiving public attention. He thinks up a plan to leap down from the roof top of the Empire State building in full view of everyone. Assuming that he was able to accomplish the jump unhurt, can we not still ask whether his achievement is a good for him, regardless of the fact that it fits his overall desire to receive public notice and this in turn fits in with the network of his overall plans and desires? Perhaps we might want to question not so much the goodness of the action as the rationality of the man. But this is not to say that his plan or achievement is inconsistent with the overall system of his desires. Rather, our consideration would be directed to the fact that his action violates some commonly accepted stipulations of rationality (not to endanger one's life unnecessarily).

Thus, it does not seem so obvious that the third part of Rawls' definition has clearly shown how a person's good can be defined in terms of a rational plan of life without having to appeal to some ideal of rationality. Assuming, however, that such a nonmoral definition is possible and that the remarks here are essentially mistaken, the question yet remains: precisely how does a person determine whether a particular life plan in any given range of life plans is indeed a rational one?

To answer this question, we need to examine the two conditions specified by Rawls for determining the rationality of a life plan.

... a person's plan of life is rational if, and only if, (1) it is one of the plans that is consistent with the principles of rational choice when these are applied to all the relevant features of his situation, and (2) it is that plan among those meeting this condition which would be chosen by him with deliberative rationality, that is, with full awareness of the relevant facts and a careful consideration of the consequences. (17)

Two key notions emerge in these remarks: the notion of the principles of rational choice and the notion of deliberative rationality. For a plan of life to be deemed rational, its choice must be motivated by the principles of rational choice and by deliberative rationality. When these two conditions have been met, only then can one say that a particular plan of life is rational or not. Also, on the basis of these conditions a distinction could be made between an "apparently" rational plan and a "really" rational plan. Our next task therefore would be to give some content to these two notions. What does Rawls mean by the principles of rational choice and the concept of deliberative rationality?

17 Rawls, op. cit., p.408.

THE PRINCIPLES OF RATIONAL CHOICE.

Rawls adopts the method of enumeration in an attempt to articulate the principles of rational choice. He adopts this method on the assumption that full enumeration of the principles would necessarily unveil the content of rationality. He begins his enumeration with the more familiar and least controversial aspects of rationality.

1. Principles of effective means: (Given) a particular objective that is wanted, and that all the alternatives are means to achieve it, while they in other respects remain neutral, one is to adopt the alternative which achieves that objective with the least expenditure of means or given the means, one is to fulfill the objective to the fullest possible extent.
2. Principles of inclusiveness: One short term plan is to be preferred to another if its execution would achieve all of the desired aims of the other plan and one or more further aims in addition.
3. Principles of greater likelihood: (Where) the aims which may be achieved by two plans are roughly the same. Then it may happen that some objectives have a greater chance of being realized by one plan than the other, yet at the same time none of the remaining aims are less likely to be attained. A greater likelihood of success favors a plan just as the more inclusive end does. (18)

The most important criticism that might be brought against these principles -- the counting principles as Rawls calls them -- is that they do not provide adequate basis for ordering plans in accordance with their rationality and hence

18 Ibid, pp. 411-3.

their goodness. They do not apply in all situations. For, as Rawls himself points out, "the means may not be neutral, inclusive plans may not exist and the objectives may not be sufficiently similar"⁽¹⁹⁾. Moreover, the principles merely provide a strategy for counting existing desires and wants in an attempt to see how many of these can be realized. There is no indication of a criterion for a further analysis of our desires and for judging the relative intensity of our wants. The desires and wants of individuals are simply taken⁽²⁰⁾ as given.

Rawls is aware of this inadequacy. He is also aware that an adequate account of rationality ought to yield principles which would enable an individual to determine which of his several wants or desires is rational. Thus he tries to fortify the counting principles by appealing first to "one formal principle that seems to provide a general answer"⁽²¹⁾. The latter is the principle which requires an individual "to adopt that plan which maximizes the expected net balance of satisfaction. Or ... to take that course most likely to

19 Ibid, p.417.

20 Rawls asserts in "Goodness to Fairness" that "the interests and ends of individuals depend upon existing institutions and the principles of justice they satisfy". This, he says, is in contrast with the doctrine of abstract individualism which maintains that the fundamental aims and interests of individuals are determined independently from particular social forms.

(22)
realize one's most important aims". Again, this formal principle does not seem very helpful. As Rawls notes, it "fails to provide us with an explicit procedure for making up our minds".⁽²³⁾ Even while appealing to it, the individual is yet "to decide what it is that he most wants and to judge the comparative importance of his several ends".⁽²⁴⁾

DELIBERATIVE RATIONALITY

With the apparent failure of this formal principle, Rawls goes on to consider the notion of "deliberative rationality" which he takes over from Sidgwick. He modified the notion rendering it as follows:

... the rational plan for a person is the one (among those consistent with the counting principles and other principles of rational choice once these are established) which he would choose with deliberative rationality. It is the plan that would be decided upon as the outcome of careful reflection in which the agent reviewed, in the light of all the relevant facts, what it would be like to carry out these plans and thereby ascertained the course of action that would best realize his more fundamental desires. (25)

21 Rawls, op. cit., p.416.

22 Ibid.

23 Ibid.

24 Ibid.

25 Ibid., p.417.

This definition of deliberative rationality is believed to capture in essence what Sidgwick refers to as "a person's future good on the whole". That is, what a person "would now desire and seek if the consequences of all the various courses of conduct open to him were, at the present point of time, accurately foreseen by him and adequately realized in imagination"⁽²⁶⁾. Also, Nielsen remarks that the conception of rationality emerging in the definition is "fundamentally what Richard Brandt is after in his analysis of rational desires and in trying to ascertain what is intrinsically desirable in terms of them, and it is ... a reasonable way of reading Mills' and Dewey's talk about the relation between the desired and the desirable"⁽²⁷⁾. However, apart from these veritable sources of support, the notion of deliberative rationality as expounded by Rawls does seem capable of generating some support of its own. His account does seem to provide considerable insight into what could be described as a rational desire, and what it is for an object to be rationally desired. For instance, unless someone has a pro-attitude toward a proposed course of action, there does not seem to be any good reason why he should go ahead to perform the act. And, an action can be said to be rational only if there are

26 Ibid, pp. 416-7.

27 Kai Nielsen, op. cit., p.116.

reasons for which an agent performs it. However, despite its strong appeal, Rawls' formation of the principle of deliberative rationality is still plagued by serious difficulties. Even if we are to assume that the "other principles of rational choice" mentioned in the formula could be made sufficiently determinate, and that we are able to determine, using Rawls' criteria, what the fundamental desires for an individual are, ⁽²⁸⁾ the question still remains: why is the choice of plan made by an individual under full information necessarily rational or good? Admittedly, it is not unusual for individuals to have passionate desires for certain things even when they do not have full information about these things. And, they may yet persist in their desire after they come to know all the facts about these things. Imagine two persons X and Y with contrasting attitude towards the university: X wants a university education and Y hates everything about the university. Even when the complete information about the university has been made known to each of them, they yet persist in their attitudes. Now, if on Rawls' account, we accept "full information" as the overriding factor, then we cannot

28 Rawls, as a matter of fact, does not think that these assumptions could be demonstrated conclusively. After a lengthy remark about "the devices of deliberation and time related principles", he concludes that "the determination of rational plans is indeterminate in important ways. The more evident and easily applied principles of rational choice do not specify the best plan; a great deal remains to be decided". (Rawls, op. cit., p.449.)

but see the contrasting attitude of X and Y as being equally rational. There would be no basis whatever why we should rank any of them as being more rational than the other. Thus we would be saying, in effect, that whatever anyone happens to desire would be rational provided it is desired with full information. Not only does this proposition fail to square with our ordinary intuitions about rationality, it is essentially relativistic. In the final analysis, it is no more than saying that what people happen to desire under such conditions is objectively rational, and hence good. As Hare would argue it seems possible to envisage a situation whereby a good number of people do in fact have similar desires. But an adequate theory of the good should be able, at least in principle, to show such desires to be irrational. Consider the penchant for wearing a tan in this society during the summer months. Recent researches have indicated that this practice is capable of inflicting skin cancer. We might say, of course, that a lot of people may not be aware of such hazards. But suppose they come to know and they yet persist in sun bathing? Would the fact that they now know all the facts about their action necessarily make it good?

The practical implications of Rawls' position could be far reaching. If the best plan for an individual is the one that he would adopt if he possessed full information, and if this plan is the objectively rational plan for him and

determines his real good, why would the state want to meddle with individuals who knowingly choose to take their own lives. For instance, Garry Gilmore, the convicted murderer in the state of Utah has knowingly and freely chosen to pay for his crime by facing a firing squad. Now why was it necessary for the amnesty board of the State of Utah to review his case? Accepting Rawls' formula, could the Board ever come forward with reasons to set aside Gilmore's desire to die by firing squad? Would such reasons not undermine the objectively rational plan for Gilmore and thereby jeopardize his real good? But why could they (the members of the board) not rule that: "although Gilmore has elected to die by firing squad in full awareness of the nature and consequences of his desire, we have commuted his death sentence to life imprisonment in conformity to right reason and the laws of the land". It does seem to make sense for one to argue sometimes that even though an agent wants a certain object, knowing all the facts about the object, yet the object is not the best or the most rational thing for him to want. But with Rawls' conception of deliberative rationality there is hardly a basis for such an argument.

Another important objection that could also be brought against Rawls -- Nielsen does in fact mention a similar objection -- is that it is not a conceptual anomaly for one to conceive of a thoroughly rational plan of life

which is totally evil or at least not good. (Recall the question of mobsters.) Even if we want to condemn the anti-social activities of these men as irrational, we cannot do this by appealing to the Rawlsian conception of rationality. It might indeed be the case that some men become members of the underworld without full information as to what all the implications are. Some people might even be drawn into it against their will. But it is also a known fact that several people have become members with full "deliberative rationality". That is, they know all the facts that could be known in regard to this way of life. Are we then to conclude that since this latter group of men adopt their plan of life with full information, their plan is objectively rational, it is the best plan for them, and one which determines their real good? It is not so obvious how Rawls can overcome such a paradox.

However, there is a passage in Rawls that tends to invalidate the foregoing objection. In drawing a parallel between utilitarianism and his own contract doctrine, Rawls indicates that the indeterminacy of rational plans "is no difficulty for justice as fairness, since the details of plans do not affect in any way what is right or just"⁽²⁹⁾. Since individual ways of life must conform to the principles of justice that are arrived at independently, the arbitrary features in

29 Ibid, p.449.

individual plans become irrelevant. Thus Rawls concludes that the indeterminacy in the notion of rationality "does not translate itself into legitimate claims that men can impose on one another. The priority of the right prevents this". Rawls gives an interesting example in support of his point. He says:

(30)
... assume that the larger part of society has an abhorrence for certain religious or sexual practices, and regards them as an abomination. This feeling is so intense that it is not enough that these practices be kept from the public view; the very thought that these things are going on is enough to arouse the majority to anger and hatred. Even when these attitudes are not supportable on moral grounds, there appears to be no sure way to exclude them as irrational. (31)

It would seem from this example that Rawls would, after all, allow the possibility whereby the desires of a person could be rational and yet be morally untenable. By the same token, there could be immoral life plans that are thoroughly rational (at least in the non-moral conception of rational). However, Rawls cannot consistently allow these possibilities in view of some of the remarks he makes in his discussion of the question "Why Be Moral?". He contends that, at least under the circumstances of a well-ordered society, "being a good person is indeed a good". In such a society,

30 Ibid.

31 Ibid, p.450.

32 Rawls, p.398.

he argues, and perhaps more generally as well, having a sense of justice is a good. Hence, having moral attitudes is "desirable from the standpoint of rational persons who have them when they assess their situation independently from the constraints of justice".⁽³³⁾ The implication here is that there are good and sufficient reasons independently of moral considerations that could motivate a reflective person to be a moral person with genuine moral attitudes.

It will become evident when we begin to probe the congruence argument, that Rawls cannot consistently maintain these claims, let alone justify them. This fact notwithstanding, there are several remarks made by Rawls towards the end of his book that would seem to support an earlier suggestion that there might be immoral life plans that are yet rational. Recall a citation in the last chapter where Rawls admits the possibility of some individuals not having sufficient reasons (as defined by the thin theory) to affirm their sense of justice. Since he does not specifically refer to these individuals as being irrational, we may presume that they are at least capable of performing acts which are rational even if such acts go against the constraints of morality.

Is one to conclude then, that the foregoing interpretation of Rawls' notion of deliberative rationality is inaccurate?

33 Ibid, p. 399.

This conclusion is not so obvious. On the contrary, the assumptions about rationality in Rawls' example are, as Nielsen rightly observes, far more crudely Humean than those he usually makes and hardly fit with his conception of deliberative rationality.⁽³⁴⁾ For instance, if the dominant attitude in the hypothetical society is not one which is fully informed but rests merely on intensity of feeling, then it is obviously inconsistent with what Rawls himself describes as deliberative rationality. (A desire is rational only if it is based on full information.) However, it might yet be the case that the attitude may help promote a certain life plan which is rational. Should those religious or sexual practices pose a threat to the rational life plans of these agents, the agents can hardly be considered rational if they do not rebuff them. But even when the latter is the case, the attitude of the agents is no less arbitrary insofar as it lacks sufficient information. In fact, the agents are appealing to their prejudices as distinct from their moral beliefs.

The objections raised so far seem to suggest a fundamental weakness in the notion of rationality underlying the thin theory of the good. As was mentioned earlier, Rawls, in an attempt to make his original position appear morally neutral, adopts the notion of rationality that is standardly used in bourgeois economics (rationality that is purely instrumental). Thus,

34 Nielsen, op. cit., p.119.

rationality is identified with the maximization of individual utility. Rawls accepts this identification subject to two constraints: (1) that there is a class of primary goods such that an increase in these goods always represents an increase in utility; and (2) that his rational contractors are free from envy. But does such a conception of rationality adequately describe what is ordinarily considered rational in non-moral discourse? Also, could such a conception ever lead to the notion of morality articulated by Rawls in justice as fairness (morality adhered to for its own sake).

Perhaps, it should be emphasized that Rawls' account of rationality cannot be rejected out of hand as incorrect. There is no doubt that the account does capture some important ways in which the term "rational" functions in ordinary (non-moral) discourse. For instance when an individual is ordinarily asked to be rational, this might just be a way of telling him to pursue that course of action that is capable of maximizing his interests. Also, when individuals come together to enact a contract, there is a tacit understanding that each party is out to achieve benefits which otherwise, would have been unattainable, or at least more difficult to attain, without such a collective instrument. Therefore, the difficulty with Rawls' account does not lie in identifying the maximization of individual expectable utilities as a conception of rationality. Rather, it lies in an attempt to treat this conception as if it were the concept of rationality.

There are numerous instances in which a person can act rationally even when he has every reason to believe that by so acting he would not or cannot possibly be maximizing his own personal gains. For example, several Israeli soldiers were known to have formed themselves into bridges over barbed-wire fences to enable their troops to infiltrate into Lebanon during the 1967 war. Although their action was cited as heroic, it can hardly be said that they did it in order to maximize their individual gains. It might not even be necessary for a person to want (have an affective attitude) to perform an act, before the act can be considered rational. Insofar as he has sufficient reasons for performing the act, when he acts on those reasons and with no apparent overriding countervailing reasons, then his action can properly be regarded as rational.

Some clarification might still be necessary here, as the foregoing claim is well open to challenge. What is meant by "sufficient reasons" here has not been explained. Someone might respond that crazy and indeed insane people can and do have reasons. Consider the case of paranoids and their insane rationales. Sometimes we try to use principles of rational choice or a conception of a rational person to help us distinguish good reasons and rare (unusual) rationales from bad reasons or insane rationales. Nevertheless, on Rawls' account something can only be called a reason for an action if it accords with a particular life plan which is rational. And, to determine whether a life plan is rational, all we need

is to find out what a person would want or desire under conditions of full information and vivid recall. To this, one could only respond that it is possible, at least theoretically, for the agents in Rawls' example to continue to show aversion to those sexual and religious practices even when they must have had "full information" about these practices. And given Rawls' account of rationality, there is no basis for regarding them as irrational persons. If this is indeed what Rawls' theory of rationality leads to, it shows a definite weakness in it.

Apart from the inadequacy of the received view of rationality as an approximation of the concept of rationality, its effectiveness in generating the concept of morality is highly suspect. In an elaborate and very sophisticated examination of the notion of rational co-operation, David Gauthier concludes among other things that the received view of rationality can only lead to the "morality of economic men". (35)

In such a morality, the relations of men one with another are

accidental, external; it is no part of their nature, as they conceive it, that each should affect the well-being of others in whatever ways he does. Hence, to them morality appears as artificial and instrumental, as a human contrivance whose only rationale is to assure the greater well-being of each. If as moral men they treat each other as ends, accepting for others the same relative advantage they demand for themselves, yet in a deeper sense they treat each other as means, for the

35 David Gauthier, Nous, 8 (1974), p.62

principles of optimization and fairness are to each but part of an overall policy of individual utility-maximization". (36)

It is not very certain whether Rawls can effectively respond to Gauthier's criticism. Although he seems quite anxious to reconcile morality with rationality -- the entire third part of A Theory of Justice is devoted to the task of relating a theory of justice to the theory of goodness as rationality -- it is not so evident whether the morally neutral conception of rationality he invokes can accomplish this task. The most that Rawls can hope to get from this conception is the idea of a collective morality. And if it is indeed the case that the morality of his men does not transcend the morality of rational co-operation, then the possibility of a morality of principles (morality for its own sake) emerging from A Theory of Justice can almost be ruled out.

Certain aspects of Rawls' conception of rationality will be reviewed again in subsequent chapters. Also, some of the general claims of the received view will be examined in Chapter Eight. In that chapter, an attempt will be made to show how an enriched conception of rationality would be more consistent with the kind of morality Rawls is trying to characterize and explain. Meanwhile, the discussion on the thin theory will conclude with a brief examination of what Rawls designates as

the Aristotelian principle. The Aristotelian principle is presented by Rawls as one of the conditions of rational choice.

SECTION 11

THE ARISTOTELIAN PRINCIPLE

The formulation of what Rawls refers to as the Aristotelian principle is another important attempt by him to impose substantive constraints on the possible patterns of human motivation. Rawls admits the fact that his three-stage definition of the good is "purely formal". According to this definition, the good for an individual consists in choosing that plan which will maximize his expectable utilities from any maximal class of plans. Even though Rawls goes on to specify a network of principles from which the notion of deliberative rationality and the principles of rational choice derive, "We still cannot", he admits, "derive from the definition of rational plans alone what sorts of ends these plans are likely to encourage". (37) To be able to draw conclusions about these ends therefore, "it is necessary to take note of certain general facts." (38) Accordingly, Rawls postulates

37 Rawls, op. cit., p. 424.

38 Ibid.

a basic principle of motivation which he refers to as the Aristotelian principle. This principle states that:

... Other things being equal, human beings enjoy the exercise of their realized capacities (their innate or trained abilities), and as this enjoyment increases the more the capacity is realized, or the greater its complexity. (39)

Rawls provides some clarification to this statement:

The intuitive idea here is that human beings take more pleasure in doing something as they become more proficient at it, and of two activities they do equally well, they prefer the one calling on a larger repertoire of more intricate and subtle discrimination. For example, chess is a more complicated and subtle game than checkers, and algebra is more intricate than elementary arithmetic. Thus the principle says that someone who can do both generally prefers playing chess to playing checkers, and that he would rather study algebra than arithmetic. (40)

The Aristotelian principle is said to be a deep psychological principle of motivation:

It accounts for many of our major desires, and explains why we prefer to do some things and not others by constantly exerting an influence over the flow of our activity ... expresses a psychological law governing changes in the pattern of our desires. Thus the principle implies that as a person's capacities increase over time (brought about by physiological and biological maturation, for example, the development of the nervous system in a young child), and as he trains these capacities and learns how to exercise them, he will in due course come to prefer the more complex activities that he can now engage in which calls upon his newly realized abilities. (41)

39 Ibid, p.426.

40 Ibid, p.426.

41 Ibid, pp.426-8.

And in the theory of the good:

(The Aristotelian principle) states a deep psychological fact which, in conjunction with other general facts and the conception of a rational plan, accounts for our considered judgments of value. (42)

The most obvious problem that emerges from these statements is that of interpretation. How does one interpret the Aristotelian principle? In which category of beliefs is it to be located? Brian Barry sees only two options: either "it is an empirical generalization or it is a partially constitutive definition of 'rationality' (43)." But neither of these two options, he thinks, seems to augur well for Rawls' theory. Let us examine the two interpretations.

First, let us suppose that Rawls intends the principle as part of a substantive definition of rationality. There are two passages that one could cite in Rawls' book to support this assumption. In one instance Rawls says that, since "the Aristotelian principle is a feature of human desires as they now exist, rational plans must take it into account". (44) - Also he declares in another context that the "things that are commonly thought of as human goods should turn out to be the ends and activities that have a major place in rational plans. This principle is part of the background that regulates these

42 Ibid, p.432.

43 Brian Barry, The Liberal Theory of Justice, p.24.

44 Ibid, p.432.

judgments." (45) Thus, we might want to say that the Aristotelian principle is an essential and unique feature in human nature if "rationality must take it into account". It is more than just an accidental feature in human nature. In the second passage the principle emerges much stronger as one of the criteria for assessing or evaluating human goods. It is one of the ground rules that regulate our practical judgments concerning human goods. For our judgments of human goods to become intelligible, we need to take the Aristotelian principle into account.

Therefore, if the Aristotelian principle is interpreted as part of the definition of rationality, it would then follow that any choice or behaviour pattern that is not in accordance with it must be considered irrational. Hence, for example, let us suppose there is an individual who likes to solve puzzles for his leisure. If he does not generally aspire to attempt the more difficult ones, then he must be deemed irrational. Indeed some legitimate questions might be raised about the character traits of such an individual even if one does not go so far as calling him irrational. Did J. S. Mill not insist that it is better to be a Socrates dissatisfied than a pig satisfied?

45 Ibid.

However, Rawls cannot consistently maintain such a strong interpretation of the Aristotelian principle because it involves an appeal to some ideals of human excellence. Rawls explicitly rejects such an appeal in his critical remarks about the ethics of perfectionism. He argues that perfectionism is inadequate as a theory of morality because we do not have a sure way of determining what the ideals of human excellence are. An alternative reading must therefore be sought for the Aristotelian principle.

The second interpretation suggested by Barry appears to be a viable alternative. The Aristotelian principle could be read as a statement from empirical generalization. Rawls himself seems to be leaning in this direction when he refers to the principle as a "deep psychological fact". If the Aristotelian principle must rank as a psychological fact, then, like other psychological facts, its statement must derive from empirical generalization.

There are several passages in Rawls' book that tend to support the latter interpretation. In what he refers to as an "evolutionary explanation" of the Aristotelian principle, Rawls says that:

Natural selection must have favored creatures of whom this principle is true. Aristotle says that men desire to know. Presumably we have acquired this desire by a natural development, and indeed, if the principle is sound, a desire to engage in more complex and demanding activities of any kind as long as they are within our reach. (46)

But Rawls does not construe this evolutionary explanation in a justificatory light:

The evolutionary explanation, even if it is correct, is not of course a justification for this aspect of our nature. In fact, the question of justification does not arise. (47)

The latter passage appears rather puzzling in view of the one cited two pages earlier. It seems difficult to think of the Aristotelian principle as "part of the background that regulates" our judgments about those things that are commonly thought of as human goods, and yet the notion of justification does not enter into it. Does this imply then that the question of rational justification is irrelevant to our judgments concerning human goods? Perhaps what Rawls has in mind here is the idea that regulatory principles are not themselves susceptible to direct justification within the system they support. But should this indeed be his thinking does this not take us back to the first interpretation of the Aristotelian principle as part of the criterion of what it is to be rational?

Perhaps there is some satisfactory way of resolving this issue. At any rate, the problem itself is not very significant at this stage of our discussion of the Aristotelian principle. Rather, it seems quite evident from the two

46 Ibid, p.431.

47 Ibid, p.432.

passages cited above that the Aristotelian principle can be correctly described as an empirical generalization. One of Rawls' examples would seem to lend further support to this claim:

Thus imagine someone whose only pleasure is to count blades of grass in various geometrically shaped areas such as park squares and well-trimmed lawns. He is otherwise intelligent and actually possesses unusual skills, since he manages to survive by solving difficult mathematical problems for a fee. The definition of the good forces us to admit that the good for this man is indeed counting blades of grass, or more accurately, his good is determined by a plan that gives an especially prominent place to this activity. (48)

The point of this "fanciful case", Rawls says, is only to show that the correctness of the definition of a person's good in terms of the rational plan for him does not require the truth of the Aristotelian principle. The definition would still be satisfactory even if the principle should prove inaccurate or fail altogether. (49) Thus, it emerges that the statement of the Aristotelian principle falls into the category of those propositions that can be checked for their truth or falsity. Should this then be the case, that is, should the Aristotelian principle be read as an empirical generalization, it would seem necessary, contrary to Rawls' remark in the passage just cited, that its proposition be true

48 Ibid, p.432.

49 Ibid, p.433.

enough for it to function effectively as a premise in the original position. The contractors cannot be made to reason from false premises. This is one of the considerations that lead Rawls to modify some of the assumptions at the original position in A Theory of Justice. In his earlier paper "Justice as fairness" Rawls stipulates that the parties in the original position are to choose their principles of justice on the assumption that their enemies were entirely in control of their future. He modified this constraint in A Theory of Justice since it entails reasoning from false premises. Also in his "Goodness to Fairness" Rawls states quite emphatically that the premises of the thin theory must be sound and correct. He says:

So far as possible, one assumes only a knowledge of the thin theory of the good, a theory founded on the facts of psychology and setting out the general structures of human desires and ends. On the basis of this theory, we can define the primary goods. Of course, the thin theory must be sound. (50)

It appears therefore that Rawls must actually rely on the correctness of the Aristotelian principle if the principle is to yield all what he expects to get from it. But precisely how does one establish the truth or correctness of this principle? It might be the case that most people most of the time do, in fact, derive more interest in attempting

50 Rawls, "Goodness to Fairness", op. cit., p. 538.

more challenging tasks. But even such a considerably weaker claim is seriously challenged by Barry. He insists that the Aristotelian principle is "as a matter of fact false for most people most of the time"⁽⁵¹⁾. Let us suppose that Barry is wrong (though this is not so obviously the case). Yet, can it be said that everyone's central goal in life is to realize his most complex and innermost abilities? How do we account for those hundreds and thousands of students who drop out of school every year in this country just "to take things easy"? We also have the Gandhies, the Moonites and the hippies who cannot be made to see the ethic of developing or actualizing their natural talents or complex abilities. Of course, we need to be mindful of the constraint, "all things being equal". The Aristotelian principle does not claim that people will, in every circumstance and under any situation, have a higher-order desire to realize their most complex and innermost abilities. Rather it says that people will have this desire, if there are no intervening factors to frustrate it. Hence, since the parties in the original position have been stripped of the various contingencies that might threaten this motivation, they will take it into account in their choice of principles of justice. But how satisfactory is this response? Is the phrase "all things being equal" any more than a platitude?

51 Brian Barry, op. cit., p.29.

When is it that "all things are ever equal" in real life? Perhaps one might say that this is only a practical difficulty and not a theoretical one. While this is true, it is also important to note that the principles of justice deriving from the Aristotelian principle are to apply in a world in which all are never equal.

It does not seem that the Aristotelian principle can be shown to be true or correct if it is interpreted as a strong empirical law as Barry's second alternative seems to indicate.

However, it might not be necessary to approach it in such a strong sense. Why should the principle not be interpreted as a general tendency law? Thus, while it might yet be seen as an empirical generalization, it need not be true of every individual human being living on this planet. Rather, it might just be a general tendency that is true of most people most of the time.

Such a weaker rendering of the Aristotelian principle would seem to have two significant merits. It is quite consistent with a fallibilist position: a position that Rawls could more easily identify with. Recall one of his remarks cited earlier. If, as he says, some individuals may not have sufficient reasons (as defined by the thin theory) to affirm their sense of justice, it would also be the case that some individuals may not have the desire to act

in accordance with the Aristotelian principle. Also a weaker interpretation of the Aristotelian principle would certainly dismiss the problem of a conclusive justification. In fact, this would make the principle look much more like a psychological principle. (Psychological laws are by nature, high-power assumptions; they are not susceptible to absolute or conclusive justification.)

However there are two important reasons why this interpretation of the Aristotelian principle may yet fall through. First, if the principle is merely stating a general tendency, why does Rawls attach so much importance to it? Why is it so necessary for the rational persons in the original position to take it into account in deliberating on principles of justice. We must remember here that Rawls cannot appeal to any ideals of rationality or human goodness. Such an appeal will defeat the very purpose of the thin theory. Apart from the fact that such a move would go against the claim that the conditions of the original position are a constrained minimum, it entails an appeal to the very ideals which the neutral assumptions of the thin theory want to yield. Hence the most Rawls can say is that the Aristotelian principle is a unique feature about humans. (It is one of the "general facts" about humans.) But to this, we could quickly respond that the tendency to laugh is also a unique feature about humans. Only human beings have the tendency and the capacity to laugh. Now

why is it more rational (rational in a non-moral sense) for the parties in the original position to take the tendency to act in accordance with the Aristotelian principle into consideration and not the tendency to laugh? It is not so evident how this question can be answered satisfactorily without an appeal in the direction of human excellence.

The second question raises a difficulty similar to the one mentioned towards the end of the last chapter. Let us suppose that Rawls indeed wants to take the fallibilist route. In other words, we assume that he wants to assert the Aristotelian principle as a general tendency among humans (a desire that is true of most people most of the time). Let us also assume that the unique importance of the principle can also be demonstrated without appealing to some ideal (though I do not see how this could be done). The question still remains: is Rawls not in effect using a desire that is only true of most people most of the time as a basis for generating principles of justice which are to apply equally both to those who have this desire and those who do not have it? Again it does not do for Rawls to respond that since most people in society have this desire, and since those who do not have it are the odd ones (most people in society are not like these) it was alright for him to invoke the principle. For such an argument to be successful, Rawls would want to show that only those individuals whose behaviour

accords with the Aristotelian principle are rational. He would have to show that only those individuals who have this higher-order desire can truly reflect the "general facts" of human psychology, and that others who do not have the desire are deviants from the "general facts" of human psychology. In other words, he must show that anyone who does not act in accordance with the Aristotelian principle is less than rational. Again, Rawls cannot pull such an argument through under the constraints of neutrality of the thin theory. Perhaps such an argument might be possible under normal conditions. In fact, an appeal to some standard such as the Aristotelian principle seems quite implicit in several of our everyday evaluations. For instance, when a person's performance is evaluated, very often the evaluation is done against the background of what he is capable of achieving or what he ought to achieve.

However, Rawls cannot proceed in this direction because such judgments are not value-free. For him the task is to show from the standpoint of a thin conception of rationality (a value-free conception of rationality), that anyone who does not have the higher-order desire stated in the Aristotelian principle is less than rational or less of a human being. And unless he is able to show this, it is difficult to see how he can escape the charge of deriving the principles of justice as fairness from assumptions which are

not true of all rational persons. He would need to show that justice as fairness, unlike teleological theories, is not merely appealing to a dominant conception of the good.

Rawls is apparently aware of this difficulty and he tries to deal with it in "Goodness to Fairness". While conceding the fact that the original position as a whole might not be neutral between conceptions of the good, he argues that it is still as fair as possible to the parties conceived as moral persons. He says:

And, of course, the OP idea as a whole is not neutral between conceptions of the good in the sense that the principles of justice adopted treat them all equally. Any definite agreement is bound to favor some views over others. The most that can be expected, then, is that the OP is fair to the parties (conceived as moral persons). Even this may be asking too much, since there may be no very clear notion of fairness for this case, or at any rate, no notion of perfect fairness. But it is still worth asking whether the OP is as fair as possible; or whether some other account of the initial situation would be more fair. In asking this question, it is sometimes helpful to view the matter as follows: we want to work out principles of accommodation between different moralities much as a constitution insuring liberty of conscience and freedom of thought contains principles of accommodation between different religions. Some may think, just as many have held in the religious case, that no principles of accommodation between moralities exist; or else that there is no fair initial situation in which such principles, granting they were available, could be agreed to, even hypothetically. Perhaps so. Still we would like to know why the OP is not fair and how far it falls short. (52)

Perhaps, it should be mentioned once more that Rawls can hardly be indicted on the charge of "Platonism" as alleged by Nowell-Smith in view of his remarks in this passage. He does not intend to capture the idea of perfect fairness; nor does he intend the construction of his original position to be entirely free of moral constraints. Also, it would seem that if the observations in this passage are well taken, at least some of the important objections raised both in this chapter as well as in the previous one would have been effectively countered.

However, there are several reasons why the argument in this passage cannot work. First, if Rawls indeed believes that the original position cannot be neutral between everyone's conception of the good, then, not only is the claim about the preference for the primary goods (that the primary goods are the things that every rational person must want, if he wants anything at all) false, but his attempt to establish the congruence between justice and goodness would also be unnecessary. The thrust of the congruence argument is not to show that the good of a morally good person is congruent with his sense of justice. If this is the proposition that Rawls wants to show, not only will the entire exercise be trivial, he would not be addressing himself in any significant manner to the "Why Be Moral?" question. All he would be saying is that those who have already attained the sentiments of justice

can only attain their good by acting justly. On the contrary, what raises interest in the congruence argument is his attempt to show that it is desirable to have moral "attitude from the standpoint of rational persons who have them when they assess their situation independently from the constraints of justice".⁽⁵³⁾ But the claim being made here is quite different from the one cited earlier. What Rawls seems to be saying here is that there are good and sufficient reasons independently of the constraints of morality, which rational persons are capable of grasping and which could motivate them to want to be morally good persons. Hence, the modification suggested by Rawls in the argument being reviewed cannot work because it is inconsistent with one of the major features in the program of justice as fairness. It does not fit the spirit and purpose of the congruence argument.

Another major reason why Rawls' argument does not succeed is that we do not have a sure way of carrying out the comparisons it demands. As it was argued in the first Chapter, the successful use of the original position as a critical device requires, among other things, the validity of the assumption that one is capable of knowing all the possible premises that could lead to a particular conception of justice. To rank one conception over another,

53 Rawls, A Theory of Justice, p.139.

it must be the case that one is able to know all the possible premises that could lead to each of these conceptions and be able to determine which conception yields the best possible sets of premises. Without having this knowledge, which of course is practically impossible, any rejection of one conception in favour of another would be premature. Rawls' rejection of the principle of average utility on the basis of his own derivation of the principle is premature. Also, the question that Rawls wants us to consider in assessing the fairness of his original position is rather loaded. It assumes that rational people by virtue of an instrumental rationality must show a preference for such goods as liberty of conscience and freedom of thought. It was argued in the last chapter that even though most rational persons might want these goods most of the time, there could be legitimate reasons why a rational person might forgo them for some other goods.

Perhaps one important point that must also be taken into consideration in assessing both the thin theory as well as the entire program of the original position is the need for Rawls to be absolutely correct in the assumption that lead to his principles. If he wants to establish a principle of association for members of a well-ordered society, the outcome of such principles must be rationally acceptable to the representative members of such a society. David Gauthier was both articulate and very convincing on this point. He

outlines two different standpoints in which the derivation of Rawls' principles must be assessed. He says:

In the derivation, it is essential to focus on two different standpoints. First, of course, is the standpoint of the reasoner in the original position, choosing "behind a veil of ignorance". But, second, there is the standpoint of the reasoner after the principle is chosen, the resulting society is instituted, and the veil of ignorance is lifted. I want to insist that a principle establishing the terms of association must be rational, not only prospectively, but also retrospectively. By this I mean, not that the choice made in ignorance be the choice one would make, if one could, when one is aware of the actual circumstances governing the outcome, but that awareness of the outcome should not lead one to judge the original choice irrational, when one in imagination reimposes the constraints under which it was made. The choice must be rationally acceptable, in this sense, to each representative person after the veil of ignorance is lifted. This requirement is, it seems to me, implicit in Rawls' insistence "that the original position be interpreted so that one can at any time adopt its perspective". (54)

The significance of Gauthier's insistence becomes all the more obvious if one alludes to the fact of an assumed connection between morality and rationality. The parties in the original position are expected to make their choice of principle in such a way that they would still consider their choice the most rational they could make even when the veil of ignorance must have been lifted. But if the objections that have been raised especially in these two chapters on the thin theory are valid, then we have no reason to believe that all the

54 David Gauthier, "Justice and Natural Endowment: Toward a Critique of Rawls' Ideological Framework", op. cit., pp.9-10.

members of the original position would be equally satisfied when the veil of ignorance must have been drawn. At least, those of them who do not have preference for the primary goods, and those who do not seek the higher-order desire stipulated by the Aristotelian principle may not be satisfied with a well-ordered society which is organized in accordance with the principles deriving from these assumptions. Rawls' insistence that the original position should be fair to the parties considered as moral persons does not help the situation very much. He has not given us an adequate criterion for determining a moral person. Is it so evident that his conception of a moral person fits that of everyone else? However, this is not the problem at hand. But should it be the case that he indeed wanted the original position to focus on the notion of moral persons, it would seem from the criticisms raised in this chapter that he cannot make much headway in this direction because his thin conception of rationality is by far too thin to generate the notion of a moral person. With these remarks, we now proceed to examine Rawls' full theory of the good.

CHAPTER SIX

THE FULL THEORY OF THE GOOD

Having defined a person's good as the successful execution of a rational plan of life, and his lesser goods as parts thereof, Rawls goes on to introduce further definitions. What he tries to do in the latter definitions is to relate the theory of justice to a theory of goodness as rationality. Rawls refers to this aspect of his theory as the full theory of the good. He needs the full theory for a definition of the moral worth of persons and for a definition of beneficent and supererogatory Acts.⁽¹⁾ Also, he wants to be able to determine on the basis of the full theory whether or not "being a good person is a good thing for that person".⁽²⁾ Here, by appealing to "the principles of moral psychology", Rawls tries to show how the sense of justice would ideally be acquired by members of a well-ordered society. He then goes to argue that a well-ordered society would allow for individual autonomy, encourage the realization of community, diminish the inclination to envy and spite, and enable its members to express their nature as free and equal moral persons.

1 Rawls, A Theory of Justice, pp.438-9.

2 Ibid, p.397.

Thus in contrast to the thin theory, the full theory of the good takes into account the two principles of justice as fairness as well as the formal constraints of the concept of right as developed from the original position. With these principles of right and justice on hand, the full theory of goodness as rationality can then be applied without restrictions to the issues mentioned in the preceding paragraph: issues which, according to Rawls, "have an important place in moral philosophy"⁽³⁾.

Thus an attempt will be made in the next two chapters to probe the success of the full theory of goodness as rationality as it applies to the two cases of a morally good person and a good society. In this chapter we will examine Rawls' definition of a good person and the psychological process of developing an effective sense of justice. In the next chapter, we will examine the congruence of goodness and justice.

A GOOD PERSON OR A PERSON OF MORAL WORTH

Rawls does not think that his account of the "good" would raise any serious difficulties. Many philosophers, he claims, would be quite willing "to accept some variant of goodness as rationality for artifacts and roles, and for such normal values as friendship and affection, the pursuit of knowledge and the enjoyment of beauty, and the like"⁽⁴⁾. This

3 Ibid, p.433.

4 Ibid, p.434.

acceptance is due to the fact that the main elements of goodness as rationality are, as Rawls contends, "extremely common ... (they are) shared by philosophers of markedly different persuasions"⁽⁵⁾. However, when it comes to the question of moral worth, this conception of the good does not seem adequate since it is believed to express an instrumental or economic theory of value. It is generally assumed, as Rawls notes, that when people speak of the just or benevolent person as being morally good, the notion of goodness involved is different from that expressed in goodness as rationality. However, Rawls maintains that the full theory of goodness as rationality can in fact account for these judgments once the principles of right and justice are on hand. He says:

The reason why the so-called instrumental or economic theory fails is that what is in effect the thin theory is applied directly to the problem of moral worth. What we must do instead is to use this theory only as a part of the description of the original position from which the principles of right and justice are derived. We can then apply the full theory of the good without restriction. (6)

Rawls considers three possible ways of making such a transition.

First of all, we might identify some basic role or position, say that of citizen, and then say that a good person is one who has to a higher degree than the average, the properties which it is rational

5 Ibid.

6 Ibid.

for citizens to want in one another. Here the relevant point of view is that of citizen judging other citizens in the same role. Second, the notion of a good person could be interpreted as requiring general or average assessment so that a good person is one who performs well in his various roles, especially those that are considered more important. Finally, there may exist properties which it is rational to want in persons when they are viewed with respect to almost any of their social roles. Let us say that such properties, if they exist, are broadly based. (7)

The third approach is, according to Rawls, the most plausible one. Not only can it be made to include the first as a special case, it does capture the intuitive idea of the second. Now as we focus on this approach the first difficulty that confronts us is how to identify the point of view from which the broadly based properties are rationally preferred and the assumptions underlying such a preference.

Rawls tries to resolve this difficulty. First he asserts that "the fundamental moral virtues, that is, the strong and normally effective desires to act on the basic principles of right, are undoubtedly among the broadly based properties"⁽⁸⁾. The truth of this assertion, he believes, can

⁷ Ibid, p.435.

Rawls takes the notion of broadly based properties from T.M. Scanlon. He illustrates this idea as follows: "... in the case of tools, the broadly based properties are efficiency, durability, ease of maintenance, and so on. Much less broadly based properties are properties such as keeps its cutting edge, edge does not rust, and ... By analogy, a good person, in contrast to a good doctor or a good farmer, and the like, is one who has to a higher degree than the average person the broadly based properties (yet to be specified) that it is rational for persons to want in one another". (Rawls, op. cit., p.435.)

hardly be questioned if it is supposed "that we are considering a well-ordered society, or one in a state of near justice"⁽⁹⁾. Rawls indeed takes this setting to be the case. Thus he argues that "since the basic structure of such a society is just, and these arrangements are stable with respect to the society's public conception of justice, its members will in general have the appropriate sense of justice, and a desire to see their institutions affirmed"⁽¹⁰⁾. But, as Rawls further notes, "it is also true that it is rational for each person to act on the principles of justice only on the assumption that for the most part those principles are recognized and similarly acted upon by others"⁽¹¹⁾. Hence he concludes that the "representative member of a well-ordered society will find that he wants others to have basic virtues, and in particular a sense of justice. His rational plan of life is consistent with the constraints of right, and he will surely want others to acknowledge the same restrictions"⁽¹²⁾. To strengthen this assumption, "we should also like to be sure that it is rational for those belonging to a well-ordered society who

8 Ibid, p.436.

9 Ibid.

10 Ibid.

11 Ibid. (emphasis mine.)

12 Ibid.

have already acquired a sense of justice to maintain and even to strengthen this moral sentiment"⁽¹³⁾. With these presumptions on hand therefore, the fundamental virtues can be classified among the broadly based properties. And should any member of a well-ordered society possess them to a higher degree than the average member, the same shall be deemed a good person.

Before considering any objections to this formula, it might be useful first to clarify a complication that Rawls himself envisages in connection with the notion of "broadly based properties". Apart from moral virtues, there are other properties such as intelligence and imagination, strength and endurance, that may also be included in the notion of broadly based properties when applied to human beings. In fact Rawls does point out that a certain minimum of these attributes are a necessary condition of right conduct. Unless a moral agent is guided by judgment and imagination, his "benevolent intentions may easily lead to harm"⁽¹⁴⁾. Nevertheless, it remains the case that "unless intellect and vigour are regulated by a sense of justice and obligation, they may only enhance one's capacities to override the legitimate claims of others"⁽¹⁵⁾. In the absence of moral constraints, a situation whereby "might is right" might easily prevail. But from a social point of

13 Ibid.

14 Ibid.

15 Ibid.

view, the possession of these natural assets cannot but be desirable if society is to attain its material goals. Hence they must, within limits, be included in the class of broadly based properties.

Thus Rawls insists on the necessity of distinguishing between the moral virtues and the natural assets:

The latter we may think of as natural powers developed by education and training, and often exercised in accordance with certain characteristic intellectual or other standards by reference to which they can be roughly measured. The virtues on the other hand are sentiments and habitual attitudes leading us to act on certain principles of right the virtues can be singled out by using the conception of justice already established; once this conception is understood, we can rely on it to define the moral sentiments and to mark them off from the natural assets. (16)

In the light of this distinction, a good person may then be characterized as one who has to a higher degree than the average, the broadly based features of moral character that it is rational for persons in the original position to want in one another. But why would persons in the original position want these broadly based features of moral character in one another? Rawls' answer to this question is rather straightforward:

Since the principles of justice have been chosen, and we are assuming strict compliance, each knows that in society he will want the others to have the moral sentiments that support adherence to these standards. Thus we could say alternatively that a good person has the features of moral character that it is rational for members of well-ordered society to want in their associates. (17)

The merit of this definition as Rawls sees it, is that the notion of goodness as rationality has been extended to persons without appealing to new ethical notions. All that is required is the theory of justice which has the thin account of the good as a subpart; by applying this to persons, the notion of moral worth is then accounted for. The relevant question now it seems, is the adequacy of this definition. How does it square with our ordinary conception of a morally good person?

Rawls is careful to note one initial objection that might be raised against a definition of a good person along the lines of goodness as rationality. This is, that such an approach is bound to fail since it is by no means a self-evident fact that man qua man has a definite role or function. In regard to functional objects such as knives, watches, cricket-bats or even soldiers, goodness as rationality might make a strong appeal since the term "good" is supervenient upon the properties which describe the function of these objects. But when we talk of man as such, it is not immediately

16 Ibid, p.437.

17 Ibid,

This explanation tends to suggest that moral sentiments or virtues can only be seen and prized by the members of the original position as instrumental values. Since the men have already chosen the principles of justice, they could only be considered rational, if they desire the sense of justice or any appropriate moral sentiment that would ensure the general compliance with these principles. Now how does one generate an intrinsic morality from such a standpoint?

evident what role or function we are to ascribe to him. However, Rawls insists that the full theory of the good undermines this objection, since it is not necessary to suppose "that persons have some particular role, much less that they are things to be used for some ulterior purpose"⁽¹⁸⁾. In spite of this claim, Rawls still does not offer a definition of 'good' (moral) simpliciter. Rather, he makes many assumptions, and, in particular, he assumes that "being a member of some community and engaging in many forms of co-operation is a condition of human life"⁽¹⁹⁾. And as members of a well-ordered society individuals have a shared final end of promoting their just institutions.

Thus it emerges that even in the case of moral goodness, the term "good" does not operate in a vacuum. It supervenes upon some natural (observable) properties: features of moral character that it is rational for persons in a well-ordered society to want in their associates. Therefore, the two important notions that we need to examine in this definition are: (1) the notion of a well-ordered society; and (2) the fundamental moral virtues, especially the virtue of justice. Both of these must be adequately accounted for if Rawls' definition of a morally good person is to succeed.

18 Ibid, p.438.

19 Ibid.

Now, assuming that there are indeed such broadly based
(20) properties or features of moral character that members of a well-ordered society might desire to see in one another, the question may yet be asked: why does an individual desire a well-ordered society at all? Perhaps this question might seem improper at first. Rawls, as noted above, already assumes that being a member of some community and engaging in many forms of co-operation is a condition of human life. Besides, one of the assumptions underlying the choice of the principles of justice is strict compliance -- each person knows that in society he will want the others to have the moral sentiments that support adherence to these standards. Thus Rawls' formula might be stated deductively as follows: if rational mutually disinterested persons are concerned to further their individual interests in a well-ordered society, then they must desire in one another such properties or features of moral character that generate a well-ordered society. And, whoever possesses such features to a higher degree than the average member of a well-ordered society, such a one is a good person.

At this point, the question raised above becomes very important. There is no doubt that what Rawls is trying to achieve is an objective characterization of the morally good such that

20 The precise status of these are yet to be determined.

whatever may be described as a moral good would be seen and appreciated by reflective persons as an intrinsic good (a good which, as Kant would say, is 'unconditionally' good and desired for its own sake). However, in the order of execution, it is not certain whether Rawls' definition so far has been able to transcend an instrumental or a utilitarian conception of the good. Indeed if an agent already desires a well-ordered society then the rationality of taking the most effective means towards a desired end, makes it objectively necessary or rational for him to desire in his associates the broadly based features of moral sentiments that make a well-ordered society possible. But is a well-ordered society within such a framework any more than a utilitarian good? Is it anymore than a means which enables an individual to realize his conception of his own good whatever such a conception might be? The point here is not whether people are, as a matter of fact, motivated by consideration of their own good in seeking to become members of some community and engaging in many forms of co-operation. Rather, the question is whether there is adequate justification for projecting such a natural tendency as a moral ideal to be sought and maintained for its own sake.

In Rawls' account a well-ordered society looks very much like what a psychiatrist would mean when he speaks of "mental health" to his client: it is that without which a

person cannot be satisfied, no matter how well placed he is in life. Hence, one might say that anyone who voluntarily rejects "sound mental health" must be irrational. But would such a conclusion be valid in every conceivable situation? Suppose the client of the psychiatrist does not, for some reason, believe that human satisfaction or happiness is anything worth striving for. Is there any conclusive reason why he should desire "good mental health"? Perhaps Gary Gilmore, the Utah State Prison murderer, would be a good example here. Gilmore was aware both of his crimes as well as the penalty he had to pay. But rather than spend the rest of his life behind the prison bars, he chose to be executed. Now, would Gilmore be so evidently irrational if he does not desire "good mental health" on the morning of his execution? Perhaps there are certain traits about Gilmore's behaviour on account of which we may conclude that he is thoroughly irrational. But such a conclusion cannot be based solely on the fact that he does not desire good mental health. Good mental health is as useless as life itself to a man who knows that he is about to be executed.

Thus one serious weakness about Rawls account is that his notion of moral goodness does not generate its own support. (21)
An appeal must be made to some other end in order to provide

21 This is one of the main objectives that Rawls wants to achieve especially in trying to maintain a rapport with Kant.

adequate justification for its desirability. And should an individual renounce this other end, there would be no conclusive reason for him to adopt the moral point of view. The fundamental virtues or moral sentiments which generate a well-ordered society are objectively necessary only when an individual has already committed himself to a well-ordered society. If he sees a well-ordered society as a necessary means for attaining his own good, only then is he obliged by rationality to desire the fundamental moral virtues. But should his good as he conceives it not require a well-ordered society, then it would be superfluous to require such an individual to possess any moral virtue. For instance, an individual who conceives his good along a career as a member of the underworld would hardly desire any sentiments of justice or a well-ordered society. None of these is relevant in the attainment of his good as a Robinhood. Again, we might want to say that it is irrational for any one to desire to be a member of the underworld. While we might be able to cite good reasons in support of this, Rawls' calculative model of rationality is not a very helpful premise. With this narrow conception of rationality -- a conception that yet functions as a subpart of the full theory -- it is not immediately evident that it is a failure in rationality should an individual voluntarily reject a well-ordered society.

Rawls might yet insist that the desire for gangsterism or any anti-social activity does not pose a problem for his account. With the constraints of right and the two principles of justice known publicly to govern the society, the desire for anti-social behaviour is effectively ruled out. Here again we seem to be talking of a person who already has a moral commitment to the ideal of a well-ordered society. But our problem is why an individual (every individual) should be morally committed to this ideal in the first place. Why must everyone who lives in a well-ordered society, and who knows that such a society is governed by a public conception of justice, desire as an individual to be genuinely just? In other words, why is it not possible for a free-rider egoist to want to be a member of Rawls' well-ordered society? Perhaps we do not really need such an extreme case as that of a free-rider egoist. But should any prudent person at all desire a well-ordered society as a means of attaining his own good, such a one may desire the fundamental moral virtues in order to be seen to be rational. But why does his desire that the society be well-ordered and governed by the principles of justice commit him as an individual to have a genuine desire for the moral virtues of justice?⁽²²⁾ Or, assuming that

22 It is important to note here that what is collectively rational for an individual to do is one thing. What is individually rational for a particular individual in a reasonably well-ordered society to do is another.

he does in fact desire these virtues in himself in order to be able to demand it of others, what value or worth do they have for him?

The foregoing raises a very crucial question: is the desire to act justly regulative of everything that is to count as a rational life plan, such that acting justly can always be seen as part of any man's good? This question naturally leads to Rawls' argument of congruence. Rawls would have us believe that the principles of justice as fairness and the notion of goodness as rationality are necessarily congruent. Thus, even if rational persons are generally inclined to desire a well-ordered society and the principles of justice, such a desire is not just a means to some arbitrary ends, but a desirable end (an end that is good in itself). Hence, the degree to which an individual desires this end would define his moral worth.

The discussion for the remainder of this chapter will focus on the congruence argument. First we will examine Rawls' account of how an individual acquires the sense of justice and the moral virtues which define his moral goodness. Then in the next chapter we shall try to determine whether or not Rawls has successfully demonstrated that being a good person is always a good thing for anybody at all. The issue to be considered here is whether Rawls' full theory of the good does offer a good criterion for discriminating between a morally good person and a person of good morals.

THE DESIRE TO ACT JUSTLY

On Rawls' account, one of the fundamental virtues that members of a well-ordered society would rationally desire to see in one another (one of the virtues that they must have in order to be considered as morally good persons), is the sense of justice or the desire to act justly. This sentiment disposes individual members to adopt and to want to act from the moral point of view to the extent that the latter (the moral point of view) is defined in part at least by the principles of justice. As such, the desire to act justly becomes regulative of anything that is to count as a rational life plan, and the sense of justice in turn becomes part of every man's good.

The idea of the sense of justice emerges as a very important concept as Rawls addresses himself to the problem of stability. He maintains that, in order to have a well-ordered society with stable institutions and morally good persons, it is necessary that individual members within the society have a shared sense of justice. Hence, for Rawls, it would not be incompatible with the notion and structure of a well-ordered society should the life plans of individuals in it vary from one person to another. Individual life plans give prominence to different aims, and persons are left free to determine their own good. But, despite the heterogeneous conception of the human good among its members, stability yet

prevails in a well-ordered society because every individual is motivated by the desire to act justly. Thus Rawls characterizes a well-ordered society as:

one designed to advance the good of its members and effectively regulated by a public conception of justice ... it is a society in which everyone accepts and knows that the others accept the same principles of justice, and the basic social institutions satisfy and are known to satisfy these principles. (23)

Two initial questions seem most pertinent at this point: (1) how does an individual come to acquire the sense of justice?; and (2) why is a person necessarily committed to the moral point of view, hence becoming a morally good person, just because he has acquired the sense of justice? These questions will provide the guideline for the discussion in this section. The focal point is Rawls' transition from facts of individual moral sentiments -- what he designates as "general facts of moral psychology" -- to universal ethical norms. Can this transition undermine the common objections to ethical naturalism?⁽²⁴⁾ For instance, if, as Rawls suggests, the liability to moral feelings seems to be as much a part of an individual's natural sentiments as his tendency to be joyful, or the liability to grief, is it yet valid to describe moral feelings as intrinsic goods: things that must be sought for their own sake? We begin our inquiry with Rawls' account of the three-stage development of a moral view.

23 Ibid, p.452.

Before embarking on his sketch of the development of a moral view, Rawls notes that "a moral view is an extremely complex structure of principles, ideals, and precepts, and involves all the elements of thought, conduct and feeling"⁽²⁵⁾. He emphasizes that "many kinds of learning ranging from reinforcement and classical conditioning to highly abstract reasoning and the refined perception of exemplars enter into its development"⁽²⁶⁾. Thus in attempting his sketch, Rawls limits himself to the special case of how the moral development might occur in a well-ordered society: a society realizing the principles of justice as fairness. His aim is to indicate "the major steps whereby a person would acquire an understanding of an attachment to the principles of justice as he grows up in this particular form of a well-ordered society"⁽²⁷⁾.

Thus Rawls distinguishes between three types of moralities: the moralities of authority, of association and of principles. His account of the development of these moralities "is to be tied throughout to the conception of justice

24 Although Rawls' account of the good differs significantly from R.B. Perry's definition of the good as "object of anyone's interest or interest in any object" it is no less a naturalistic account insofar as it takes natural observable properties of persons into account.

25 Rawls, op. cit., p.461.

26 Ibid.

27 Ibid.

which is to be learned, and therefore presupposes the plausibility if not the correctness of this theory. " (28)

THE MORALITY OF AUTHORITY

This is the first stage in the sequence of moral development. While certain elements of this form of morality carry over into adult lives this morality in its primitive form is essentially that of the child. It consists mainly in the basic rules of conduct (elementary do's and don't's) that those in authority (parents and parent surrogates) introduce to the child to help him distinguish what is right from what is wrong. Although the child is not able to assess the validity of the precepts or injunctions addressed to him, he is expected to submit to the authority of his parents if he is to act rationally. There is, of course, the possibility of his parents having biased and distorted moral and social views. This does not create any problem. In fact, this possibility is, according to Rawls, ruled out by the postulate of a well-ordered society. In a well-ordered society, "everyone accepts and knows that the others accept the same principles of justice ...". (29)

28 Ibid.

Again we might recall here an earlier remark concerning the necessity of an adequate survey of the essential features of Rawls' theory before considering his account of the "good". The plausibility of his full theory of the good presupposes

However, for the inculcation of the morality of authority to be successful, it is essential that a relationship of love and trust exist between the child and his parents or those in authority. Without this relationship, it is rather impossible for the child to gain any effective and moral understanding at all. In the absence of love and trust, the child might yet obey the precepts of his parents. But such obedience would only derive from the fact that he fears them. He senses their power over him and the only way he can save his 'neck' is by being subservient to their powerful will. In such a "love-less" situation the child does not develop a sense of what it is for there to be a moral or a de jure authority. He can only perceive authority as a means for intimidating 'others' to submission.

In order to generate the morality of authority in the child, it is necessary therefore, that his parents first manifest a genuine sense of love towards him. Their love to the child should consist, not only in caring for his needs and wants, but in affirming "his sense of the worth of his own person". Through such love, the child in time becomes

the correctness of his theory of justice. However, if the objections raised in the first part of this study are well-founded, then there is at least a prima facie case for rejecting Rawls' full theory of the good.

29 Rawls, op. cit., p.454.

30 Ibid, p.464.

"assured of his worth as a person. He becomes aware of himself as a person who is appreciated for his own sake by what are to him the imposing and powerful persons in his world". From this he will come to develop a sense of trust and confidence in both his parents and his surroundings, and also, a sense of self-esteem in himself.

The initial love of the parents would elicit the love of the child for his parents in return. But this is not to construe the love of the child in an instrumental sense:

The child's love does not have a rational instrumental explanation: he does not love them as a means to achieve his initial self-interested ends. With this aim in view he could conceivably act as if he loves them, but his doing so would not constitute a transformation of his original desires. (31)

The point of this caveat is rather obvious. Although a morality of authority is a primitive form of morality, it is yet crucial that it should possess the essential characteristic of genuine morality if it must feature as one of the major premises in Rawls' moral geometry. Above all, it ought to be able to generate its own support. Hence, a direct causal analysis of the reciprocity of the child's love cannot be adequate as this might subsequently reduce the entire morality of authority to a utilitarian or a consequentialist morality. Thus we must

31 Ibid.

32 Ibid.
The original desires of the child is to satisfy his self-interests as these appear to him.

suppose that the love of the parents is not tied to the good performance of the child. Also, the love response of the child is not motivated by his self-interest. Either way it is intrinsic love: a love for its own sake. As Rawls says, a moral view cannot develop under conditions of punishment or reward:

The child's having a morality of authority consists in his being disposed without the prospect of reward or punishment to follow certain precepts, that not only may appear to him largely arbitrary but which in no way appeal to his original inclinations. If he acquires the desire to abide by these prohibitions, it is because he sees them as addressed to him by powerful persons who have his love and trust, and who also act in conformity with them. He then concludes that they express forms of action that characterize the sort of person he should want to be. (33)

Rawls' analysis of a morality of authority does look quite reasonable. Even a Skinnerian behaviourist who does not have any use for "inner" morality may yet agree with Rawls that a love atmosphere is an ideal positive reinforcement that could easily motivate a child to learn the right conduct. An injunction issued with love is most likely to be well taken and this could foster good relations between those in authority and their subjects.

However, despite the plausible assumptions made in Rawls' analysis, it yet remains a curious question whether

33 Ibid, p.466.

his morality of authority does in fact transcend a consequentialist morality. As was mentioned earlier, even though Rawls is dealing with the primitive stage of moral development at this level, his overall program is to show that the steps whereby an individual comes to have the moral point of view reproduce the entire structural scheme of morality. In particular he wants to show that children have a morality and that an appeal to authority has a genuine, if limited, place in the moral firmament. Now, Rawls' aim is not in question. Rather, what is at issue is whether his premise does necessarily yield the type of morality he has in mind.

In ordinary circumstances, as was mentioned above, the precepts or injunctions which the parents pass on to the child might stem directly from their own personal biases or prejudices. But this possibility must be ruled out because of the constraints of a well-ordered society. In fact, a morality of authority must be subordinated to the principles of right and justice before it can be a genuine part of morality. Now, while these constraints may indeed ensure that what the child receives is some form of morality, they cannot adequately ensure a specific type of morality. They cannot rule out the possibility of the child being inculcated with a utilitarian or a consequentialist morality. For instance, the mere notion of a well-ordered society ruled by the principles of justice and right does not tell us how the

parents came to adopt the moral point of view that they communicate to the child. Suppose the parents see themselves through the Hobbesian vision of the nature of man and society. Even though they would only be concerned to further their own individual interests, they might yet find it individually advantageous to subscribe to a morality and a well-ordered society ruled by justice because of man's natural conditions of misery. Now if such conditions form the basis of the parents own morality, why can they not effectively transmit such a Hobbesian morality to their child? Even if we want to say that such a morality is 'artificial' and nothing more than rational self-love, it could yet be the regulative principle of a well-ordered society.

But Rawls does insist on a non-instrumental love as a crucial factor in the development of a morality of authority. The parents love the child for his own sake and not for his good performance. They manifest their love not only by being concerned for his wants and needs, but in affirming his sense of the worth of his own person. And, on the part of the child, he too returns his parents' love not for the prospect of reward or punishment but because he sees his parents as powerful persons who have his love and trust. Perhaps such a conception of love might indeed be effective in the geometry of a non-consequentialist morality. Nevertheless, it is doubtful whether such a love is common to all of humanity. It is not even

necessary that all rational persons should possess every aspect of it. Perhaps in an individualistic oriented society, such as the North American society, a sense of individual worth and uniqueness can be rightly considered as one of the necessary virtues for personal survival. Hence, it might be a natural thing for parents in this society to want to inculcate such a sense in the child. But in a rather homogeneous society where the accent is on the community and not the individual, the sense of a person's own individual worth can hardly be considered a significant virtue. In fact, in several oriental societies and Communist China, in particular, the sense of an individual's own worth can only be seen as one of the major obstacles that a person must overcome in order to attain a meaningful existence. In such societies, as it is in many African societies, parents see themselves only as part of the community and their children as an extension of themselves. A well reared child in this setting is not one who stands out as an "individual" but one who fits in as part of the community. Thus it would seem that Rawls' assertion that "to love another means ... to affirm his sense of the worth of his own person" does not capture what the concept of love means in every culture. The affirmation of a person's sense of worth in himself is not always implied in the concept of love. In fact, in certain instances, to love someone is just to be concerned for his needs and wants, and in some restrictive sense, to be one with that person.

Along with the problem of Rawls' characterization of love, his insistence that the child's love does not have a rational instrumental explanation does present some difficulties. Perhaps when the child comes of age and becomes fully aware of "the sense of his own worth as a person", his love for his parents might take a new dimension. He might come to love his parents for what they are: powerful and loving companions. But Rawls does not provide any significant explanation why the child would reciprocate with such a 'metaphysical' love at the very initial stage. If one is to keep strictly within the confines of a psychological explanation it seems more plausible to characterize the love of the child as essentially instrumental. The child sees all the rewards and affections he gets for being well behaved and the prospects of punishment if he misbehaves. Thus when he manifests his love for his parents -- this consists mainly in behaving well -- he does so for the reward that his love can generate. What operates here is only a question of cause and effect. It is difficult to introduce non-instrumental explanation without an appeal to some intuition. Besides, the very notion of "child" seems to be in danger of being unduly stretched by the highly abstract love which Rawls postulates.

It would seem from these remarks that the most important difficulty that Rawls is bound to encounter in his analysis of moral development is the transition from

ordinary psychological facts to the moral realm: especially showing that the sense of justice is an intrinsic moral virtue. This question would be examined in some detail after reviewing Rawls' analysis of the two other stages of moral development. Meanwhile, we turn to the second stage which Rawls calls "the morality of association".

THE MORALITY OF ASSOCIATION

This is essentially a morality of a person's station and duties. Its content is given by "the moral standards appropriate to the individual's role in the various associations to which he belongs."⁽³⁴⁾ Thus the morality of association is characteristically heterogeneous since there are innumerable roles and associations that an individual might wish to, and does in fact, enter in any society. However, the ideals governing these varied associations yet retain the force of morality insofar as they are subordinate to the principles of justice in a well-ordered society. Also, while the notion of association covers such institutions as the family, school, professional bodies, neighbourhood, sports' associations and even the community as a whole, it is assumed that everyone in a well-ordered society will have the role of citizen and there

34 Ibid, p.467.

will be a full equality of citizenship, with everyone having political views concerning the common good. Therefore, a well-ordered society would have a morality of association whereby, "the members of society view one another as equals, as friends and associates, joined together in a system of co-operation known to be for the advantage of all and governed by a common conception of justice."⁽³⁵⁾ Now one important question that one might ask here is the relation between this morality (the morality of association) and the morality of authority. How, for instance, does an individual move from the primitive morality of authority to the more complex morality of association?

In many instances a person just happens to find himself in various associations in which each member has certain rights and duties. The family is a typical example. Before a child even comes of age, he finds himself a member of a family with the right to participate in whatever privileges or assets the family might have. And as he develops the appropriate capacity, he also assumes his share in the duties of the family. A social structure such as this seems a necessary condition for the very existence of society. Now with regard to a well-ordered society, several of the precepts which constitute the morality of authority teach the child to be a good member or to do the proper thing in the various associations which he

35 Ibid, p.472.

might subsequently enter. Through such precepts the parents teach the child what it is to be a good son, student, citizen, sportsman, neighbour etc.

As a person moves through the various stations with its attendant duties, his moral understanding undergoes a corresponding growth. Eventually he will come to possess certain ideals appropriate to those roles. And as his understanding of himself and his society increases, he will begin to move in the direction of having a morality of principles -- a movement that ushers in the final phase of the moral development. This last phase is attained when a person begins to work out "a conception of a whole system of co-operation that defines the association and the ends which it serves." ⁽³⁶⁾ However, this final and all-embracing stage of the moral development can only be possible when a person is capable of understanding the roles of others and see things from their own point of view. Thus a complete development of morality necessarily transcends the morality of authority. To be morally matured, an agent must take the morality of other people's motives and intentions into ⁽³⁷⁾ account.

36 Ibid, p.468.

37 Ibid, p.469.

The relation between a morality of authority and that of association is also concretely exemplified by the factor of friendship and mutual trust. For there to be a functioning cluster of associations even in a well-ordered society, there must be relations of friendship and mutual trust. There must be an atmosphere whereby people can rely on one another to carry out their respective duties. A relation of reciprocity must develop naturally. Thus like a child who just finds his parents loving him, a person who enters a well-ordered society finds that his fellow citizens are well disposed toward him and willing to "live up to their duties and obligations" in whatever association he might choose to enter. In turn (and like the child too) he develops feelings of trust and confidence towards them. He begins to acquire attachments to them and a desire to live up to the ideals of the associations. When "others act with evident intention to affirm our well-being and at the same time, they exhibit qualities and ways of doing things that appeal to us", they arouse in us "the desire to model ourselves after them."⁽³⁸⁾ A morality of association is thus "characterized by the co-operative virtues: those of justice and fairness, fidelity and trust, integrity and impartiality."⁽³⁹⁾ The typical vices against this morality are

38 Ibid, p.472.

39 Ibid.

"graspingness and unfairness, dishonesty and deceit, prejudice
(40)
and bias."

THE MORALITY OF PRINCIPLES

This is the point Omega of the moral development. In the earlier phases of the moral development, a person's attachment to the principles of morality does not follow from considerations concerning the intrinsic worth of these principles in themselves. For instance, in the more complex forms of the morality of association, the individual is said to have an understanding of the principles of justice as well as an attachment to many particular individuals and communities. But his subsequent disposition to follow the moral standards that apply to him in his various positions is not motivated by a concern for these standards in themselves. Rather, his motive for complying with them "springs largely from his ties of friendship and fellow feeling for others, and his concern for
(41)
the approbation of the wider society."

With the morality of principles, on the other hand, the motivation of the individual is to act for the sake of the moral principles themselves. When this stage of moral development is attained, the individual becomes attached to the

40 Ibid.

41 Ibid., p.473.

highest order principles of morality themselves, so that just as during the earlier phase of the morality of association, he may want to be a good sport, say, he now wishes to be a just person. The conception of acting justly, and of advancing just institutions, comes to have for him an attraction analogous to that possessed before by subordinate ideals. (42)

The highest order moral principles become intrinsically desirable and the individual begins to see them as reasons and motives for action. His moral attitudes are no longer connected solely to the well-being and approval of particular individuals and groups, but are shaped by a conception of right chosen irrespective of these contingencies.

This is not to say that the individual acts arbitrarily when he acts for the sake of the principles of morality in themselves. His desire to act justly does not spring from a vacuum. Nor is it "a form of blind obedience to arbitrary principles unrelated to rational aims".⁽⁴³⁾ In fact, in its "normal form of right and justice, the morality of principles includes the virtues of the morality of association and authority. It defines the last stage at which all the subordinate ideals are finally understood and organized into a coherent system by suitable general principles".⁽⁴⁵⁾ When

42 Ibid.

43 Ibid, p.476.

44 Ibid, p.478.

45 Ibid.

this final stage is attained, the individual is said to be morally mature.

Now, assuming that the foregoing characterization adequately reflects what it is for a person to have a morality of principles, there is still the task of showing how a person actually attains this final level of moral development. How, for instance, does a person come simply to desire, for its own sake, to promote just institutions or to act in accordance with the principles of justice?

In dealing with this question Rawls introduces a third psychological principle. When a person has obtained a mastery of the principles of morality and understands the values they secure and the way in which they are to everyone's advantage, he is then lead to an acceptance of these principles by a third psychological law.

This law states that once the attitudes of love and trust, and of friendly feelings and mutual confidence, have been generated in accordance with the two preceding psychological laws, ... then the recognition that we and those for whom we care are the beneficiaries of an established and enduring, just institution tends to engender in us the corresponding sense of justice. We develop a desire to apply and to act upon the principles of justice once we realize how social arrangements answering to them have promoted our good and that of those with whom we are affiliated. In due course, we come to appreciate the ideals of just human co-operation. (46)

46 Ibid, pp.573-4.

The two preceding psychological laws explain the motivation of the moralities of authority and association respectively.
"First Law: given that family institutions are just, and that

It seems rather obvious from this remark that having a morality of principles and the desire to act from such a point of view need not be considered in isolation from the rational aims and interests of the individuals in a well-ordered society. Provided a person has cultivated the attitudes of love, mutual trust and friendship, then the very recognition that he and those from whom he cares are the beneficiaries of established and enduring just institutions would produce and continue to reinforce in him the corresponding sense of justice. Thus one might say that it is a sufficient condition that person should have these psychological attitudes and should live in a well-ordered society. He will come to have a sense of justice. And with the sense of justice comes the corresponding desire to apply and to act upon the principles of justice. Having attained the sense of justice, a person becomes overwhelmed with a feeling of guilt and remorse if he fails to live up to the principles of justice. Living and acting for the sake of these principles thus become the criteria for determining a person's maturity.

the parents love the child, and manifestly express their love by caring for his good, then the child recognizing their evident love of him, comes to love them. Second Law: given that a person's capacity for fellow feelings has been realized by acquiring attachments in accordance with the first law, and given that a social arrangement is just and publicly known by all to be just, then this person develops ties of friendly feeling, and trust towards others in the association as they with evident intentions comply with their duties and obligations, and live up to the ideals of their station". (Rawls. p.490)

Now that the final stage of the moral development has been laid bare, we are in position to assess the intrinsic worth of the fundamental virtue of the sense of justice: a virtue which defines a person's moral worth when conjoined with other virtues that rational persons would desire in one another in a well-ordered society. Perhaps the most crucial question that must be considered here is the nature of this virtue as it emerges in Rawls' account. To be more specific, is the sense of justice any more than a mere psychological trait (a brute fact) about us human beings, or is it a sentiment that we can rationally develop and which is desirable to have from the standpoint of morality. If the former is the case, that is, human beings when placed in the appropriate environment and having the appropriate psychological attitudes, just come to have the sense of justice: If this is the case, has Rawls provided any argument why such a psychological trait must be sought and maintained as a moral ideal? If, on the contrary, the sense of justice is to be construed as a trait that one develops through a rational thought process, has Rawls been able to show why it should not be seen as an instrumental or utilitarian good? Before examining these questions, I shall first do something to make somewhat clear the force and importance of Rawls' account of the moral development.

The desire generated by these two laws leads to an acceptance of these principles by a third law which in turn motivates the individual to desire them for their own sake.

Although some of Rawls' assumptions are rather questionable, his basic line of approach to moral development has a strong appeal. In particular, his first and second psychological laws adequately portray the sort of motivations that would dispose individuals to what he designates as the morality of authority and association. For instance, we recall how, as children, we learn the difference between right and wrong from our parents or guardians. It all began with such elementary precepts as "It is wrong to kill, to steal, and to tell lies. It is right to help others in need, to obey society's laws, and to tell the truth". Even though these precepts may have appeared very arbitrary to us (at least before we were able to reason on our own) we were attached to them nonetheless because of the way we saw our parents. We saw our parents as the most powerful and the most loving persons in the world. Hence whatever they tell us must be right. Also, when we began to interact with the world beyond our homes, there were 'moral' constraints that we grew attached to in the course of time. When we enrolled in the hockey club in our neighbourhood for example, we encountered not only the rules of the game of hockey but the rules of the club as well. We must be well disposed towards other club members, and also, we must handle the properties of the club with care. After we have spent several weeks in the club, we developed friendly ties toward our peers and began to see the rules of the club as the main guarantees of everyone's

privileges in the club. And as we joined larger and more complex associations our moral consciousness became increasingly expanded and we began to think not just about ourselves but about others too. Thus it appears that the basic idea in Rawls' analysis does effectively capture the process whereby individuals are introduced to the realm of morality and become subsequently attached to moral principles.

However, one might yet raise a few questions about the institutional setting which Rawls assumes in the formulation of the psychological laws. He notes himself that "the most striking feature of these laws (tendencies) is that their formulation refers to an institutional setting as being just, and in the last two, as being publicly known to be such".⁽⁴⁷⁾ The need for these constraints is rather obvious. Since Rawls is concerned with principles of moral psychology and not principles of psychology per se, his hypotheses can legitimately appeal to moral notions. He says:

The principles of moral psychology have a place for a conception of justice; and different formulations of these principles result when different conceptions are used. Thus, some view of justice enters into the explanation of hypotheses about this psychological process incorporate moral notions even if these are understood only as part of the psychological theory. (48)

47 Ibid, p.491.

48 Ibid.

The reasoning here seems rather straightforward as Rawls himself points out. To avoid the naturalism of the R. B. Perry type -- deriving moral principles from non-moral facts -- then, somehow, Rawls must introduce some moral notions somewhere in his premises. Perhaps a utilitarian theorist might have preferred such concepts as sympathy instead of a direct moral notion like the notion of justice. But here again as Rawls points out, different formulations of the principles of moral psychology would result when different conceptions of justice are used. Since he is interested in the notion of justice as fairness he must make the sort of assumption that would lead to such a conception.

However, while conceding some initial plausibility to Rawls' account the precise nature or value of the virtue of the sense of justice is still something of a puzzle. How does the sense of justice operate within the realm of moral values? Or, what precisely is its role? This question seems rather significant since the sense of justice is a vital premise,

49 Rawls indicates two approaches to social theory in which an appeal to moral notions might be considered unnecessary. But neither of these is quite satisfactory. The first is to try the behaviourist approach by explaining "the formation of affective ties by laws referring to the frequency of interaction among those engaged on some common task, or to the regularity with which some persons take the initiative or exercise authoritative guidance." Accordingly, the first law might state that "among equals co-operating together, where equality is defined by the accepted rules the more often individuals interact with one another, the more likely it is that friendly feelings develop between them." The second law may be seen as stating

not only in the definition of the moral worth of an individual, but in the argument for the stability of justice as fairness. Rawls wants to argue ultimately that justice as fairness and goodness as rationality are congruent. That is, given the circumstances of a well-ordered society, a person's rational plan of life always support and affirm his sense of justice. (50) For such an argument to even begin at all, it is necessary to have a clear conception of what the sense of justice is.

As seen earlier, Rawls maintains that once a person has the proper psychological attitudes and lives in a society with just institutions, he will come to have a sense of justice. He will have the desire to do his part in maintaining just arrangements, even when doing so may not be beneficial to people he happens to care about, including himself. The need

that "the more someone in a position of authority uses his powers and leads those subject to him the more they come to respect him." (Rawls p.492). While both of those laws apparently avoid the use of moral notions, neither of them mentions the justice (or fairness) of the arrangement in question, they "are bound to be very limited in scope." Besides, as Rawls argues, those who are "subject to another exercising authority will surely regard him differently depending upon whether the whole arrangement is just and well designed to advance what they take to be their legitimate interest." To ensure that there is reciprocal friendly feelings and co-operation among all the members of an association, both those in authority and the ordinary members of the association must seek to advance the legitimate interest of everyone. And this is precisely what the notions of justice in Rawls' psychological laws are trying to guarantee.

The second alternative is for one to take the route of economic theory of democracy. The latter is the "view that extends the basic ideas and methods of price theory to the

for this moral constraint is quite evident. At the level of the morality of association, a person comes to recognize how social arrangements in accordance with the principles of justice have promoted his own good and the good of those with whom he is associated. This consideration tends to engender in him "a desire to apply and to act upon principles of justice"⁽⁵¹⁾. But in order that he might be considered truly moral (moral, at least, in a non-utilitarian sense), his desire should extend beyond caring about his own welfare or the welfare of those with whom he has particular ties.⁽⁵²⁾ While friendship and the ties of association may indeed be necessary to morality, they are not sufficient conditions. As Rawls rightly observes "while every citizen is a friend to some citizen, no

political process" (Rawls 492). While this approach obviously excludes moral notions, it must, Rawls warns "for all its merits be regarded with caution" (Rawls pp.492-3). Whereas it might indeed be feasible in the realm of economic theory, "a theory of a constitutional regime cannot take the rules as given, nor simply assume that they will be followed" (Rawls p.493). In economic theory such factors as the behaviour of sellers and buyers in accordance with rules and constraints are taken as given, and at a certain level there is no objection to this. In the political process, on the contrary, even when everything is done in accordance with constitutional procedures we need to explain why these are accepted. Thus the notion of morality is a crucial element in a social theory. "The leading political actors are guided ... in part by what they regard as morally permissible, and since no system of constitutional checks and balances succeeds in setting up an invisible hand that can be relied upon to guide the process to a just outcome, a public sense of justice is to some degree necessary" (Rawls, Ibid.).

50 Ibid, p.513.

51 Ibid, p.474.

citizen is a friend to all." But in order to be a morally good person, one must act as if every citizen in the world were to be a friend.

Again, while Rawls' definition of what is to count as the sense of justice seems quite reasonable, his actual account of how we come to develop this sense is very inadequate. His account cannot provide an adequate basis for this definition. Rawls appears at least to be relying mainly on associationist psychological principles. He suggests that once the attitudes of love, trust, and friendship have been cultivated, then the recognition that we and those for whom we care are the beneficiaries of established and enduring just institutions tend to produce and continue to reinforce in us the corresponding sense of justice. Perhaps we need not demand an explanation of why this is necessarily the case, since such an explanation if possible at all, might be the proper task of psychology and not moral philosophy. But assuming that Rawls is correct in this observation, would the sense of justice (the desire to promote just institutions or to act in accordance with the principles of justice for its own sake) be anymore than a brute fact about humans? And should this be the case, why should the sense of justice be sought as a moral ideal (an intrinsic

52 For example his family, tribes or circle of friends.

53 Rawls, op. cit., p.474.

moral value to be sought and maintained for its own sake)? It is important to note here that even though Rawls makes an appeal to contingent facts, and parts of his theory are thoroughly descriptive, the overall aim of his theory is to provide a unified perspective from which we can adjudicate our differences. He seeks to provide an Archimedian point from which we can judge or evaluate social institutions as to their morality. Within such a normative framework, it is not sufficient for us just to want to be moral. Rather, we must desire to be moral in order to be seen to be rational. But has Rawls provided any reasons at all for the latter? What he has given may be considered as causes (a causal analysis of) why people generally tend to be moral. But such causes cannot undermine a consistent amoralist's challenge for a justification (and not an explanation) of the moral point of view.

However, there is a passage in which Rawls tries to provide some explanation as to why individuals should come to have the desire to act from a conception of right and justice.

How is it possible that moral principles can engage our affections? In justice as Fairness there are several answers to this questions. First of all... moral principles are bound to have a certain content. Since they are chosen by rational persons to adjudicate competing claims they define agreed ways of advancing human interests. Institutions and actions are appraised from the standpoint of securing these ends; and therefore, pointless principles, for example , that one is not to look up at the sky on

Tuesday, are rejected as burdensome and irrational constraints,... But secondly, it is also the case that the sense of justice is continuous with the love of mankind.... Benevolence is at a loss when the many objects of its love oppose one another. The principles of justice are needed to guide it.... Finally, the Kantian interpretation of these principles shows that by acting upon them men express their nature as free and equal rational beings. Since doing this belongs to their good, the sense of justice aims at their well-being even more directly. It supports those arrangements that enables everyone to express his common nature. Indeed, without a common or overlapping sense of justice, civic friendship cannot exist. The desire to act justly is not, then, a form of blind obedience to arbitrary principles unrelated to rational aims. (54)

Again, Rawls' insistence that moral principles ought to have some specific content is hardly objectionable. Consider some of the difficulties inherent in the doctrine of the purely conscientious act. As Rawls notes, the proponents of this doctrine claim that "the highest moral motive is the desire to do what is right and just, simply because it is right and just, no other description being appropriate."⁽⁵⁵⁾ There are at least two fundamental objections to this doctrine: (1) it cannot provide any positive guidance to a moral agent in search of the morally right thing to do; and (2) it is almost impossible to adduce any good reason why any and every moral agent should do the morally right thing. If no description whatever is appropriate to the desire to do what is right and just, it would seem that an agent can only rely on a special "moral sense" or

54 Ibid, p.476.

55 Ibid, p.477.

intuition to know that a particular desire, of all his system of desires, is the desire to do what is right and just. The very existence of such a mysterious "moral sense" raises a serious doubt. However, let us assume for the sake of continuing the argument that such a sense does in fact exist, and that individuals can rely on it to know what is the morally right thing to do. It is still impossible to give good reasons why it is desirable to do the morally right thing since "no other description could be appropriate" to the desire to do the morally right thing. Thus, Rawls' rejection of the doctrine of the purely conscientious act as being irrational seems to be well founded. Also, his attempt to interpret the sense of justice in "a natural way" does make a strong initial appeal.

However, when one considers the rather stringent way in which Rawls conceives of morality, it is not so evident how his "natural" approach could succeed. For instance, his explanation has not adequately shown why rational individuals must necessarily prize morality for its own sake. If, as the first point of his explanation suggests, rational persons in a well-ordered society develop attachment to the principles of morality in view of their worth as effective principles of adjudication, this fact only goes to show the instrumental value of these principles. It shows that rational persons would always want to use these principles to the extent that they effectively secure the human interests which the members of a well-ordered society seek to attain through rational

co-operation. But this does not entail the desire to prize these principles as intrinsic goods or ends in themselves. It does not imply that a rational person would always want to act on these principles himself, even when doing so is not to his advantage and when he could safely get away with not acting on them. The point at issue here is that, whereas rational persons would generally tend to act on the principles of justice (morality) because these principles advance their collective interests, we cannot proceed to assert from this that every rational person must actually desire to act on these principles himself even when acting on them would clearly be in conflict with his personal interest and he can safely avoid acting on them. For example, a rational person might desire that members of his society should pay income taxes in order to support the public institutions in the society. However, it is not obviously irrational on the part of such a rational person should he desire not to pay taxes himself. Rawls has at best, only provided an explanation why a rational individual will want the principles of morality especially the principles of justice as fairness to be the principles of adjudication between people. His explanation does not show that a rational person will always want to use these principles himself rather than just appearing to use them when it is in his interest to do so. That is to say, a rational individual will never want to be seen to be unfair, and he will see the

overall value of principled behaviour in society. He will see why rational individuals will want people to act in accordance with the principles of justice. But it is not so clear that a rational man must actually want to be fair, when it is not in his interest to be so, as distinct from wanting simply to be seen to be fair.

Rawls might insist here that being fair is what reason or rationality requires of all rational individuals. Perhaps this is so, and if it is, then both the apparent and the actual desire to be just must necessarily coincide. Hence, being just or being morally good or fair would be intrinsically desirable. In which case, being just is something all rational persons should want to do for its own sake. But Rawls has not provided adequate basis for such a strong claim. Surely the first point of his explanation does not provide a basis for it.

The second point in Rawls' explanation is the claim that the sense of justice is continuous with the love of mankind. Without the principles of justice playing a regulatory role, benevolence is at a loss when many objects of its love oppose one another. The very existence of benevolence is therefore contingent upon the principles of justice. The crux of this argument is the assumption that rational persons must love mankind. If this is correct, then rational persons must love the principles of justice since these principles are intimately connected with the love of mankind. Although the attachment of

the principles of justice within this formula could still be instrumental, we need not worry about this point. Rather, the crucial question is why rational persons would consider it natural to love all of mankind. On the other hand, Rawls has been able to show quite plausibly how an individual develops love attachments to his parents as well as associates. In a very careful sense, we might say that it is a failure in rationality if an individual does not reciprocate love feelings toward his parents and associates in a well-ordered society. But how this love gets extended beyond one's family and associates has not been adequately accounted for. Again, the need for the extension itself is rather obvious. Rawls sees the sentiments of love as playing a very important role in the psychology of moral development. While love relations would ordinarily develop out of common interests among particular individuals, (thus making love more of a consequential or relational property) the constraint of universality in the notion of morality compels Rawls to seek a superior type of love: a sort of an unconditional love which requires an individual to love not only his friends but his enemies as well. While such a non-utilitarian love may indeed be useful in the deduction of a deontological principle of morality, Rawls has not provided any convincing argument as to how an individual can come to possess such love. Perhaps one might want to say again that such an argument or explanation is outside the scope of moral philosophy. It might be the case that the proper place to

look for such an explanation would be in social psychology. But the least that Rawls could still do is to provide good grounds why one might consider it a failure in rationality should an individual not extend his love and concern beyond his close associates. He has not done the latter either.

The third and final argument in Rawls' analysis portrays the Kantian influence on him. By endorsing and desiring to act in accordance with the principles of justice, rational persons express their nature as free and equal rational beings. The desire to act justly belongs to their good as it aims more directly at their well-being. The question whether or not the desire to act justly necessarily belongs to every rational person's good is the central issue to be considered in the next chapter. We will concentrate at this point on the assumption that a person must desire to act justly in order to be a free and rational being. Are there any good reasons for accepting this assertion? Why, for instance, is an amoralist or a thoroughly unprincipled individual necessarily irrational? The sort of consideration invoked earlier might be relevant here. In ordinary circumstances, it might be the case that an individual who is rational, and who desires to be seen as such, would not want to be seen to be through and through unprincipled or immoral. But suppose he is able to keep his immorality or unprincipled behaviour a well-guarded secret, is he still irrational and unfree? Of course, if we are reasoning solely from the

constraints of Rawls' original position, we might say that to adopt the two principles of justice as fairness is just the rational thing to do. But in an actual society (whether well-ordered or not so well-ordered) an individual need not desire to be just or find it always in his rational self-interest to be just, especially when he is considering his rational life plans.

Rawls' belief along with Kant's is that an individual who refuses to act in accordance with the principles of justice or from the moral point of view is necessarily contradicting his rational or common human nature. Hence, such an individual must be irrational. Perhaps this is so, although the idea of someone "contradicting his rational or common human nature" would require a thorough de-mythologization. But has Rawls provided any persuasive reasons why it should be so? What Rawls refers to as the Kantian interpretation of this account becomes relevant at this point. The main thrust of this argument is as follows:

One may say, then, that a person who lacks a sense of justice, and who would never act as justice requires except as self-interest and expedience prompt, not only is without ties of friendship, affection, and mutual trust, but is incapable of experiencing resentment and indignation. He lacks certain natural attitudes and moral feelings of a particularly elementary kind. Put another way, one who lacks a sense of justice lacks certain fundamental attitudes and capacities included under the notion of humanity. Now the moral feelings are admittedly unpleasant, in some extended sense of unpleasantness; but there is no way for us to avoid a liability to them without disfiguring ourselves. This liability is the price of love and trust,

of friendship and affection, and of a devotion to institutions and traditions from which we have benefited, and which serve the general interest of mankind. (56)

This is Rawls' approximation of the Kantian motif why a man's common human nature commits him to the principles of justice and the moral point of view. Again we are confronted with something similar to the argument for the primary goods. For any rational individual to want anything at all, he must want such natural attitudes as friendship, love, affection, and mutual trust. If anything must count as a rational desire at all, then wanting these things is a rational desire. And whoever possess them, such a one can be said to possess rational attitudes. But, for an individual to desire these things at all, he must have a prior commitment to the principles of justice. He must genuinely desire to act on the principles of justice, as distinct from merely giving the outward appearance of desiring to act on them. Hence, wanting to act from the moral point of view is intrinsic to whatever might count as a rational desire. As such it is intrinsic to a man's nature to be moral.

One initial question that comes to mind is the accuracy of Rawls' interpretation of Kant. Would Kant appeal to considerations as these in establishing the rationale for the moral

56 Ibid, pp.488-9.

point of view? It is very doubtful. However, it might be useful to assess the merits of the content of the argument itself before considering the question of its Kantian heritage.

To an amoralist or a free-rider egoist, the first bone of contention in Rawls' argument might be his use of the words "humanity" and "disfiguring". Rawls seems to be using these words in a normative sense, and so they may be seen to be part of the moral phenomenon itself which he is expected to argue for. Why, the amoralist might ask, should an individual not disfigure humanity if indeed there is a humanity capable of being disfigured? Why should everyone commit himself to the well-being of humanity? Perhaps Rawls need not develop a separate argument in answer to this question. It suffices to note that it is because everyone, including the amoralist himself, is unavoidably involved with the natural attitudes which Rawls mentions in the passage that makes the moral point of view mandatory on everyone. Since these attitudes are internal to our very being or essence as humans, the conditions of fulfilling them creates an imperative that no person can rationally set aside. Thus, it is of our very essence to be moral.

While this sort of response, which looks indeed Kantian (at least in form), may both unsettle the amoralist objection as well as project the core point in Rawls' argument more sharply, it is very doubtful whether he can actually proceed in this direction. The facts of his argument can hardly stand

up to such a strong claim. What he has tried to show is that an individual cannot be a free-rider egoist and yet possess these natural attitudes. He has not argued convincingly that if a person prizes friendship, love, trust, and the like, he must of necessity desire the sense of justice and be willing to act on the principles of right and justice. An individual might simply adopt "the familial point of view" or "the tribal point of view". Thus he might just limit his concern for people quite deliberately to his own family and relatives or class of friends. In such a situation, the individual may still experience friendship, love and trust. In fact, a basis for such a conjecture can be found in Rawls' own account. Under what he describes as the morality of authority, a child does experience these natural attitudes without a conscious commitment to the principles of justice. The child does not have to know that there are such things as principles of justice. Having experienced an unconditional love and friendship from his parents, he too reciprocates with similar attitudes. Thus someone might be committed to "the familial point of view" while he rejects the principles of justice as firmly as the amoralist or the free-rider egoist. Such a person can still experience the natural attitudes. The principles of justice or morality do not impose a necessary condition for experiencing these attitudes.

Rawls can still counter this objection. As he argues in section 86 of his book, once we acknowledge that there are principles of morality which are collectively rational, and also that it is in the interest of each member of the society that everyone comply with them, then one cannot consistently maintain an egoistic or the familial point of view mentioned above.

...as the contract doctrine requires, the principles of justice are public: they characterize the commonly recognized moral convictions shared by the members of a well-ordered society. (They are seen by everyone as) the best choice from the standpoint of the original position. (57)

Under these constraints anyone who wants to be a free-rider egoist, that is one who wishes "to set on a systematic course of deception and hypocrisy professing without belief, as it suits his purpose, the accepted moral views", will find the psychological cost of such a venture too high a price to pay.⁽⁵⁸⁾ He will have "to reckon with the psychological costs of taking precautions and maintaining his pose, and with the loss of spontaneity and naturalness that results."⁽⁵⁹⁾ Whereas it may be an easier burden to cope with in a typically corrupt society of our times, in a well-ordered society, it is not a

57 Ibid, p.570

58 Ibid.

59 Ibid.

comfortable situation for anyone to rationally desire.

Even if Rawls' claims are correct -- although we might yet need some hard evidence from social psychology to ascertain that they are -- the case against a free-rider egoist in a well-ordered society has not been conclusively dismissed. For someone who really wants to make a sport of society and its moral principles, the sort of pathological feelings that Rawls mentioned do not pose an insurmountable deterrent. If recent events in our history such as the watergate are anything to go by, it should be evident that any group of individuals can without qualms take the society for a ride and no one may be able to detect it from their external behaviour.

Of course Rawls might retort that he is only considering a well-ordered society. Even then, as long as he bases his argument on such highly subjective elements as psychological feelings, his case remains extremely vulnerable. What guarantee do we have, to begin with, that all the members of a well-ordered society are going to have identical feelings about all their institutions or their moral principles. Assuming that the circumstances of a well-ordered society are such that its members must necessarily cultivate these feelings, why is it such a burden for an individual to rationalize on them and subsequently purge them out of his system. We might recall some Biblical mythologies here. It is said that all the angels were created good in the beginning. Even though they were

supposed to be living in a perfect state, half of them rebelled against God, and turned devils over time. Now why is it so impossible for an individual in Rawls' well-ordered society to undergo such a transformation with his fellow or love feelings, subsequently drying out? And if this is a possibility, why is such a person necessarily irrational even if he is still able to put on such behaviour that would enable him to realize his rational life plan.

At this point, it might be relevant to make a brief comment on the validity of the rapport which Rawls tries to maintain with Kant. The claim that an individual who rejects the principles of right and justice would be denying his rational human nature is ostensibly Kantian. Kant is very articulate on this point in the Ethical Doctrine of Elements. He says:

Man cannot renounce his personality so long as he is subject of duty, hence so long as he lives; and that he should have the moral title to withdraw from all obligation, i.e., freely to act as if he needed no moral title for his action, is a contradiction. (60)

While Rawls may share this belief with Kant, his method of establishing it in Justice as fairness is in direct conflict with Kant. In the preface to the Doctrine of Virtues, Kant totally rejects any attempt to justify moral principles by a direct appeal to feelings (pathological or otherwise). For him, an adequate justification of any moral claim must refer to metaphysics.

60 Kant, Doctrine of Virtue, trans., Mary Gregor, p.85.

... no moral principle is based, as people sometimes suppose, on any feeling whatsoever. Any such principle is really an obscure thought metaphysics, which is inherent in the structure of every man's reason The way the teacher talks about duty (his technique) need not always be metaphysical nor his terms scholastic, unless he wants to educate his pupil as a philosopher. But the thought must go all the way back to the elements of metaphysics, for without these the doctrine of virtue can have neither certitude nor purity nor even motive power. If we depart from this principle and begin with pathological or pure aesthetic or even moral feeling (the subjectively-practical rather than the objective) -- if, in other words, we try to determine duties by beginning with the matter of the will, the end, rather than with the form of the will, (376) as our basis -- then we shall indeed have no metaphysical first principles of the doctrine of virtue; for feeling, no matter by what it is aroused always belong to the order of nature -- But then the doctrine of virtue will be corrupted at its very source ... (61)

Perhaps we need not think of Rawls as being directly involved in the sort of moral thinking which Kant is decrying in this passage. Although such an element as the Aristotelian Principle which plays a very important role in the derivation of the two principles of justice is no more than a principle of feeling or desire, we cannot conclude outright that Rawls bases the principles of justice as fairness solely on feeling. Such constraints as the principles of rational choice and the notion of primary goods (especially the good of self-respect) are all essential premises in the derivation of these principles. And, in his Sketch of the various psychological phases

61 Ibid, p.33.

leading to the cultivation of the sense of justice, Rawls is careful to impose an institutional setting which is both just and publicly known to be such. However, while these constraints tend to suggest that Rawls does not intend to base his moral principles entirely on feeling, one is at a loss to find the type of metaphysics which Kant emphasises in the passage cited above. Although Rawls does make several references to specific themes and notions in Kant to support some crucial arguments, his remarks about these themes are rather casual, and, in no place in the book does one find the sort of hard core metaphysics which makes these notions distinctively Kantian. Of course, this fact by itself need not be seen as a weakness in Rawls' account. In fact, it might be argued that his account could be successful without appealing to any of the "murky" metaphysics of the Kantian type. It might even be the case that with an account so plagued with so many serious difficulties, an appeal to a Kantian metaphysics could only complicate issues the more. Be it as it may, one wonders if Rawls' account is Kantian in any important sense if he would not have any part of Kant's metaphysics. Is it sufficient for him to assert the form or conclusions of Kantian arguments without committing himself to the premises from which these arguments are derived?

Again it should be emphasized that Rawls need not take the route of the Kantian metaphysics with all its inherent

difficulties in order to construct a sound theory. But, since he has already started to toil with Kantian themes and expressly acknowledging their Kantian heritage, it is only consistent that he should have some commitment to their Kantian rationale. For Kant, morality is not a matter of choice but something inherent in our noumenal nature which must be discovered. This discovery is done not via empirical psychology but via thought metaphysics.

It appears, from the objections raised so far, that Rawls has not been able to develop the notion of the sense of justice as an intrinsic virtue or something to be sought for its own sake. If anything at all, he has only been able to show that we do, as a matter of fact, have certain moral sentiments including a sense of justice, and this phenomenon could be explained within a "natural" psychological framework. Also, he has shown why the men of his original position would recognize and desire the sense of justice together with other principles of morality as means of settling their future claims. He has not shown why these men will prize the sense of justice and the principles of morality for their own sake. And in the context of his theory of the good he has not shown why it is necessarily rational to be a morally good person and not just a person of good morals.

We will now proceed to examine Rawls' argument of congruence. The difficulties emerging from this argument would lend further support to the main conclusion reached in this chapter.

CHAPTER SEVEN

THE GOOD OF THE SENSE OF JUSTICE

THE QUESTION OF CONGRUENCE

As seen from Rawls' account hitherto, the concepts of justice and goodness are linked with separate and distinct principles. Each concept establishes a point of view from which institutions, actions and plans of life can be judged. A person is said to have a sense of justice if he has "an effective desire to apply and to act from the principles of justice and so from the point of view of justice."⁽¹⁾ And, a person is defined as a morally good person, if he has "the features of moral character that it is rational for members of a well-ordered society to want in their associates."⁽²⁾ Now the question of congruence for Rawls is whether these two families of criteria fit together in such a way that it can be seen to be "rational (as defined by the thin theory of the good) for those in a well-ordered society to affirm their sense of justice as regulative of their plan of life."⁽³⁾ To state it more precisely, the question is whether a

1 Rawls, A Theory of Justice, op. cit., p.567.

2 Ibid, p.437.

3 Ibid, p.567.

This remark is rather misleading. The questions "why should we be moral?" and "why should I be moral?" are two different questions. An answer that satisfies the former need not satisfy the latter.

person's disposition to endorse and to be guided by the point of view of justice always and necessarily accords with his own good.

While admitting that congruence cannot be taken for granted even in a well-ordered society, Rawls is careful to note that what is at issue in this question is not the rationality of choosing the principles of justice as fairness in the original position. The latter has been established, and if the argument for it is sound, then we are to take it that institutions are collectively rational and they are to everyone's advantage from a well defined general perspective. By the same token, it would also be rational for each member of a well-ordered society both to expect and to urge others to support these public arrangements and to fulfill their duties and obligations.⁽⁴⁾

However, the crucial problem in the question of congruence is whether "the regulative desire to adopt the standpoint of justice belongs to a person's own good when viewed in the light of the thin theory with no restrictions on information."⁽⁵⁾ And, in the context of Goodness as rationality, the problem is to show that it is always rational for everyone (at least in a well-ordered society), to desire and to want to act

4 Ibid.

5 Ibid.

from the point of view of justice and morality. That is, once a person has known that the institutions in his society are just and that others have (and will continue to have) a sense of justice similar to his (they comply and will continue to comply with these just arrangements), such a person can only be seen to be rational (as defined by the thin theory) if he affirms his sense of justice.⁽⁶⁾

Perhaps, as Rawls notes, it is important also not to confuse this problem with that of justifying the moral point of view or the standpoint of justice to an egoist.⁽⁷⁾ Even when the actions of an egoist may be seen to coincide with the constraints of justice and morality, he cannot be said to be acting from the just man's reasons insofar as his final ends are related to himself. Thus Rawls maintains that his argument does not attempt to show that an egoist would act from a sense of justice in a well-ordered society; or that it would be in

6 Ibid, p. 568.

7 The need for this distinction is rather evident. Since the egoist makes himself or his self-interest his one and only reason for action, it does not seem useful to try to justify the moral point of view to him by appealing to the reasons why an individual would want to act from the moral point of view. Yet, by imposing a limitation such as this, does Rawls not indirectly undercut the grounds and scope of his theory? By this limitation, it would seem that what Rawls is out to show in effect is not why it is rational for every rational person to want to act from the moral point of view, but rather, why it is rational for those who already have moral sentiments to act from the moral point of view. If the latter is the case, has the Archimedian point of view for judging the whole of morality not been restricted in grounds and scope? Besides, is Rawls' attempt

an egoist's own interest to transform himself into a just man, should he find himself in a well-ordered society. Rather, Rawls' concern is to demonstrate the goodness of the settled desire to take up the standpoint of justice. He wants to assess the goodness of the desire to adopt a particular point of view, that of justice.⁽⁸⁾

Rawls makes one initial assumption before proceeding to lay out the grounds for congruence. He assumes that human actions "spring from existing desires and that these can be changed only gradually."⁽⁹⁾ Again the rationale for this assumption is pretty straightforward. When an agent performs an act, (of course when he is not role-playing) he acts according to the sort of person he is and from the wants he has now. He does not act according to some other type of person he may have been under different circumstances, or from a set of different desires he may have had if he had only chosen differently. "Regulative aims", according to Rawls, "are especially subject to this constraint."⁽¹⁰⁾

to reconcile morality with a morally-neutral conception of rationality not doomed to failure from the beginning?

8 Rawls, op. cit., p.568.
But looked at like this, is not an affirmative answer to the congruence question truistic?

9 Ibid.

10 Ibid.

Hence, it is important that an individual should decide well in advance whether to affirm his sense of justice (11) by trying to assess his situation over a fairly extensive future. It would be inconsistent for him to want to preserve a sense of justice and all that this implies (being just and fair even when it is not to his personal advantage) while, at the same time, he wants to retain the option to act unjustly in the prospect of some personal advantage. Hence Rawls reiterates that the question of congruence is restricted to "those with a certain psychology and system of desires" (12).

However, Rawls is aware that, in view of these restrictions, the question now seems susceptible to an obvious and rather trivial answer.

Supposing that someone has an effective sense of justice he will then have a regulative desire to comply with the corresponding principles. The criteria of rational choice must take this desire into account. If a person wants with deliberative rationality to act from the standpoint of justice above all else, it is rational for him so to act. (13)

11 Here we might note a fundamental difference between Kant and Rawls. Rawls sees morality essentially as a matter of choice -- a choice that is made against the background of "our situation over a fairly extensive future". If an individual can adequately project himself to the future, and is able to see that his situation both now and in the future cannot be impaired if he does not display a sense of justice, then there is no prima facie reason why he should endorse the moral point of view or a sense of justice. With Kant, on the other hand, morality is of the very essence of rationality. Only those ends which incorporate morality are rational per se. To act from the point of view of morality is for someone to realize his noumenal or real self.

The question is obviously trivial if it is considered in this form. One could simply respond to it by saying that the members of a well-ordered society cannot but desire to act justly, more than anything else, being the sorts of persons they are. So long as they have acquired a sense of justice which is truly final and effective, they are confirmed in a plan of life that leads them to preserve and to encourage this sentiment in others. Acting justly and fulfilling this desire thus becomes part of their good which they must seek if they are to be rational.

But this approach misses the central issue of congruence which, according to Rawls, is "what happens if we imagine someone to give weight to his sense of justice only to the extent that it satisfies other descriptions which connect it with reasons specified by the thin theory of the good."⁽¹⁴⁾ We are to presume that the desire to act justly "is not a final desire like that to avoid pain, misery, or apathy, or the desire to fulfill the inclusive interest."⁽¹⁵⁾ Thus we should strive to see whether a person acting from the standpoint of the thin theory would yet want to confirm his sense of justice as regulative

12 Rawls, op. cit., p.569.

13 Ibid.

14 Ibid.

15 Ibid.

of his plan of life, if he were to base his considerations solely on the descriptions of what the sense of justice is a desire for. ⁽¹⁶⁾ With these background remarks, we will now proceed to examine what Rawls describes as the grounds of congruence.

THE GROUNDS OF CONGRUENCE

The first ground of congruence refers to one of the points made towards the end of Chapter Six. It involves the connection between acting justly and the natural attitudes of individuals in a well-ordered society. As stipulated by Rawls, one of the requirements of the contract doctrine is that the principles of justice should be public. The principles must be seen as characterizing "the commonly recognized moral convictions shared by the members of a well-ordered society." ⁽¹⁷⁾ It is therefore assumed that everyone, including the egoist sees them as the best choice of principles from the standpoint of the original position. Now with everyone else having an

¹⁶ Rawls articulates these other descriptions as follows: "... a sense of justice shows itself in at least two ways. First, it leads us to accept the just institutions that apply to us and from which we and our associates have benefited. We want to do our part in maintaining these arrangements. We tend to feel guilty when we do not honor our duties and obligations, even though we are not bound to those of whom we take advantage by any ties of particular fellow feeling. It may be that they have not yet had sufficient opportunity to display an evident intention to do their share, and are not therefore, the objects of such feelings by the second law. Or, again the institutional scheme in question may be so large that particular bonds never

effective sense of justice, any individual who contemplates being a free-rider (that is, pretending to affirm a sense of justice while in fact, he is prepared to act unjustly whenever the opportunity arises to further his personal interests) cannot cope with the psychological cost of such a venture in a well-ordered society. Rawls bases this claim on an earlier assertion that there is a connection between moral and natural attitudes. (18) In view of the content of the principles of justice and in view of the laws of moral psychology, an individual would naturally desire to be fair and just to his friends and to those for whom he cares, just as he would desire to be with them sharing in their joys and sorrows. Should we assume therefore, that everyone needs these attachments, then "the policy contemplated is presumably that of acting justly only

get widely built up. In any case, the citizen body as a whole is not generally bound together by ties of fellow feeling between individuals, but by the acceptance of public principles of justice. While every citizen is a friend to some citizens, no citizen is a friend to all. But their common allegiance to justice provides a unified perspective from which they can adjudicate their differences. Secondly, a sense of justice gives rise to a willingness to work for (or at least not to oppose) the setting up of just institutions, and for the reform of existing ones when justice requires it. We desire to act on the natural duty to advance just arrangements. And this inclination goes beyond the support of those particular schemes that have affirmed our good. It seeks to extend the conception they embody to further situations for the good of the larger community." (Ibid, p.474.)

17 Ibid, p.570.

18 In his sketch of the three stages of moral development, Rawls maintains that "in the context of the authority

toward those to whom we are bound by ties of affection and fellow feelings, and of respecting ways of life to which we are devoted."⁽¹⁹⁾ But under the constraints of a well-ordered society, "these bonds extend rather widely, and include ties of institutional forms, assuming here that all three psychological laws are fully effective."⁽²⁰⁾ Hence, for these attachments to flourish in a well-ordered society, everyone must desire to act justly to all.

Perhaps it suffices to reiterate the objections which were raised in the last chapter in connection with this argument. We might concede to Rawls that some connection does in fact exist between moral sentiments and natural attitudes. Although there are numerous instances in which persons known to be friends have undermined every sentiment or code of morality

situation, the child's natural attitudes of love and trust for those in authority lead to feelings of (authority) guilt when he violates the injunctions addressed to him. Thus the absence of these moral feelings would evidence a lack of these natural ties. Also, within the framework of the morality of association, the natural attitudes of friendship and mutual trust give rise to feelings of guilt for not fulfilling the duties and obligations recognized by the group. The absence of these feelings would imply the absence of these attachments ... Now the connection between the natural attitudes and the moral sentiments may be expressed as follows: These sentiments and attitudes are both ordered families of characteristics, dispositions, and these families overlap in such a manner that the absence of certain moral feelings evidences the absence of certain natural ties. Or alternatively, the presence of certain natural attachments gives rise to a liability to certain moral emotions once the requisite moral development has taken place." (Rawls pp.486-7).

19 Rawls, op. cit., p.570.

20 Ibid.

in their dealings with one another, it is still reasonable to suggest that if there is a real bond of friendship and love between any individuals or groups, then there is some moral commitment existing between them as well. It seems ridiculous for one to say that "A is my friend, but I do not feel like acting justly toward A". Also, we might want to insist that every moral human being (whether in a well-ordered society or not) does have a capacity for these natural attachments. The definition of man as a social animal does imply the ability to have these natural attitudes. As such, Rawls seems quite correct in pointing out that "every citizen is a friend to some citizens"⁽²¹⁾.

However, while conceding these points to Rawls, it is not so evident why these natural attitudes would necessarily constrain an individual to extend his desire to act justly beyond his friends and associates. The constraints of a well-ordered society together with the three psychological laws have not shown why this should be so. The most that one could derive from these psychological laws is the ethic of acting justly toward those to whom we are bound by ties of affection and fellow feeling, and of respecting ways of life to which we are devoted. Of course, it could be responded here that individuals who have lived in an ideal environment such as a well-

21 Ibid, p.474.

ordered society where there are persistent attitudes of mutual love, mutual respect, justice and fairplay, such individuals would be conditioned to these attitudes and they would "habitually" act from them no matter who they are dealing with—whether friends or non-friends. While such an explanation does seem plausible, it would not be acceptable to Rawls for obvious reasons. It undermines the elements of rationality and free choice. One who just desires to act justly simply as a result of some environmental conditioning cannot be said to have made a free and rational choice to be just. And, the very essence of Rawls' doctrine of the social contract is the attempt to reconcile the concepts of morality with rationality and to show how the fact of prior agreement could constitute the justification of the principles of justice and morality. In the context of the congruence argument in particular, Rawls wants to demonstrate that the desire to act justly is part of an individual's own rational good.

But to justify such a strong and important claim, it is not so evident whether Rawls or anybody else for that matter could rely very much on natural attitudes. While the having of moral sentiments is often manifested by natural attitudes, the having of these natural attitudes by themselves does not entail the having of moral sentiments. And, as long as we are able to maintain a real distinction between the two, we must be extremely cautious in using one of them to explain or justify the other.

The first point of the congruence argument does not present a basis for such a justification.

Rawls does make an additional consideration in support of his argument. He argues that since we cannot in general select who is to be injured by our unfairness, we must, if we are rational, desire to act justly toward everyone in order to protect the interest of those for whom we care.

For example, if we cheat on paying our taxes, or if we find some way to avoid doing our fair share for the community, everyone is hurt, our own friends and associates along with the rest. To be sure, we might consider covertly passing on part of the gains to those whom we especially like, but this becomes a dubious and involved affair. (22)

The point of this example indicates the golden rule of prudence. People are usually cautioned to think carefully before embarking on ruthless acts. A person's friends or relatives may turn out to be the first victims of his ruthlessness. As an African adage goes, "Those who live in glass houses should not be the first to cast stones in public places". However, this example does not help Rawls very much because what it suggests is at best a practical difficulty. (23) For someone who only wants to limit his care to his friends and associates, (one who endorses the familial or tribal point of view) it is always

22 Ibid, p.571

23 Also it could be argued that what is at issue in this example is not a moral but a "prudential" matter.

possible to devise a means of ensuring that his friends and associates do not suffer as a result of his action. In fact, it may turn out that his friends and associates, depending on the circumstances, may be well off by his refusal to pay his share of taxes, provided he could devise an effective means of passing on part of his gains to them. Consider someone with a total annual income of two million dollars. If he declares all his earnings, the department of Inland Revenue may tax off more than half of this money. Such taxes may in turn be used in financing public institutions that would in turn benefit hundreds of citizens. But, let us suppose that he wishes to restrict his care and concern only to his friends and family, and let us assume that these number ten in all. These ten people would be well off, if he declares only one million dollars as his total earnings to the department of Inland Revenue while he splits the other one million with them. He does not have to tell them about his dishonesty in order not to disturb their conscience. They might, depending on how he wants to handle the situation, see themselves as lucky beneficiaries of their great and generous friend. There is ample evidence that a lot of people have tried to beat the tax system in our society by way of generous gifts to particular churches, communities and friends. It is not immediately clear why such an action would necessarily hurt one's friends and dependents, should one contemplate doing it in a well-ordered society.

It would seem, from these considerations, that Rawls' first argument of congruence is rather inconclusive. Neither the psychological cost of desiring to act unjustly nor the difficulty of selecting who is to be injured by one's unfairness can sufficiently constrain one to act justly toward all the members and institutions of a well-ordered society.

The second basic consideration in the argument of congruence follows from the Aristotelian Principle, especially its companion effect: That it is a great good for a person to participate in the life of a well-ordered society or social union. Rawls' account of goodness as rationality yields, among other things, the conclusion that "rational plans of life normally provide⁽²⁴⁾ for the development of at least some of a person's powers". But as it is a basic characteristic of human beings that "no one person can do everything that he might do; nor a fortiori can he do everything that any other person can do", it is therefore necessary, that everyone should select and plan an orderly training of those abilities and possible interests he wishes to encourage. The training and realization of these abilities is most effectively attained when different persons with similar or complimentary capacities co-operate with one another with a view to "realizing their common or matching nature"⁽²⁵⁾.

24 Rawls, op. cit., p.523.

25 Ibid.

Besides, when persons "are secure in the enjoyment of the exercise of their own powers, they are disposed to appreciate the perfections of others, especially when their several excellences have an agreed place in a form of life, the aims of which all accept." (26) Thus Rawls concludes, following Humboldt, that "it is through social union founded upon the needs and potentialities of its members that each person can participate in the total sum of the realized natural assets of the others." (27) This in turn leads to the notion of the

...community of humankind, the members of which enjoy one another's excellences and individuality elicited by free institutions, and they recognize the good of each as an element in the complete activity the whole scheme of which is consented to and give pleasure to all. (28)

But for such a community or social union to be possible at all, and in order that its members might share fully in its life, there must be a general acknowledgement of its regulative conception. This, according to Rawls, means that everyone must affirm his sentiment of justice.

To appreciate something as ours, we must have a certain allegiance to it. What binds a society's efforts into one social union is the mutual recognition and acceptance of the principles of justice; it is this general affirmation which extends the ties of identification over the whole community and permits the Aristotelian Principle to have its wider effect. (29)

26 Ibid.

27 Ibid.

28 Ibid.

The general idea in this argument seems quite plausible. A social union where there is mutual respect and rational co-operation is a vital necessity if individuals are to achieve their rational life plans or realize their felt needs and cherished abilities. A good social union would satisfy desires, simple needs and foster the wishes of its members. It seems also plausible, as Rawls suggests, that a general acknowledgement of the principles of its regulative conception is a necessary condition for the existence of any union between individuals. To describe any group of individuals as a social union at all, its members must at least manifest such natural attitudes as friendship and fellow feeling together with a desire to act justly towards one another. It sounds rather contradictory to talk of a social union in which no one cares for each other's rights and interests.

However, despite its initial appeal, this argument yet raises several questions. For instance, why should everyone desire a social union, and why does the desire for the membership of a social union necessarily commit a person to a genuine affirmation of his sense of justice? Perhaps these questions might seem improper at first sight, being the sort of questions that an egoist or amoralist might want to ask.

Rawls, we recall, has stated expressly that the question of congruence is not that of justifying being a just man to an egoist. Rather, it is the question of showing whether the regulative sense of justice is consistent with the individual good of those members of a well-ordered society who already have the sentiment. However, while this general constraint may indeed limit the sort of questions that might be asked in connection with Rawls' argument, it is still important that the rationale of the idea of a social union be adequately established. In fact, if Rawls is able to provide adequate justificatory reasons why individuals ought (non-moral ought) to desire a social union, and, also, if it could be established that a sense of justice is an effective regulatory principle of a social union, then it would be valid to conclude that everyone ought to desire the sentiment of justice. Here the deduction would be straightforward. The rationality of taking the most effective means toward a desired end would provide reasonable justification why everyone or at least the members of a well-ordered society ought to affirm their sense of justice. In which case, affirming one's sense of justice is necessarily consistent with the good of the members of a well-ordered society. Thus, the first question that we need to examine here is the conception of a social union.

Despite the fact that Rawls does say several agreeable things, his second argument of congruence founders at its most

crucial premise. His social contract doctrine cannot establish a social union, nor has he provided any persuasive reason why every rational person ought to desire such a union. This conclusion is inevitable if we recall the motivational assumptions of the contract doctrine. The assumptions made by Rawls in the original position are incompatible with the notion of a social union. A social union, as Rawls himself observes, derives from the social nature of mankind. It has its basis on the fact that "human beings have shared final ends and they value their common institutions and activities as good in themselves." In a social union as in a team, human beings "need one another as partners in ways of life that are engaged in for their own sake, and the success and enjoyment of others are necessary and complimentary to our own good."⁽³¹⁾

If these remarks are correct, as they appear to be, then Rawls cannot hope to derive a social union from his contract doctrine. The men of his original position lack the sentiments as well as the aspirations necessary for the very existence of a social union. From his characterization above, members of a social union must be presumed to have sentiments of altruism. It is the feelings of love and altruism that binds them together as a union in pursuit of their common purposes and needs. But

30 Ibid, p.522. (emphasis mine).

31 Ibid.

the men of Rawls' original position who subsequently form his well-ordered society are "rational self-interested individuals". Even though they are neither envious nor jealous of one another, they do not have the slightest inclination to work for the good of the other person. Each one is only interested in himself: how to realize his own rational life plan. Thus what Rawls designates as a well-ordered society can only be a society of rational co-operators, where individuals find it rational to pool their efforts in order that each might safeguard his own interests. It can only be a pragmatic set-up whose entire justification lies in the fact that it can better serve the interests of its members than any other social arrangement. Contrary to Rawls' expectations, such a society cannot manifest any meaningful sense of morality (at least not a Kantian notion of morality). Even when the members treat each other with respect, their gesture can only be seen as a means of protecting their individual interests. They might even desire that their society be governed by the principles of justice as fairness. But this does not mean that they attach any intrinsic value to these principles. Rather, they would value the principles of justice just as they would value any other social institution, say a good banking system or credit union. They would value the principles of justice to the extent that these play an effective role in adjudicating their competing claims.

Now, assuming that these remarks are essentially correct, that is, the society that would emerge from the conditions of Rawls' original position can only be a pragmatic in-

dividualistic society of rational co-operators. Why is it not yet rational for every reflective person to desire such a society, especially since everyone has a good prospect of realizing his cherished abilities by co-operating with others of his kind? Perhaps this could be argued up to a point as Rawls tries to do in the congruence argument. We might say that those who wish to develop their cherished abilities ought to desire a well-ordered society as the most effective means of realizing their objective. While such reasoning seems quite plausible, it is still inconclusive. In the first place, the argument only refers to those rational persons who wish to develop their abilities. Since Rawls has not told us why it is necessarily rational for a person to want to develop his abilities, we can always presume that there are rational persons who might not wish to develop their abilities. If there is such a class of rational persons, it does not appear that Rawls has provided good reasons why they should be considered irrational should they fail to desire a well-ordered society. Perhaps Rawls might counter this objection by fastening on the stipulation of the Aristotelian Principle, especially its companion effect. An appeal to the Aristotelian Principle can indeed dismantle this objection. But this is possible only if the principle is defined in a stronger sense than Rawls would have it. For instance if Rawls is prepared to affirm that anyone whose choice runs counter to the Aristotelian Principle is irrational, then by the same token, one can argue that

a person who does not desire a well-ordered society is irrational. But Rawls would not make such claims because they imply a substantive ideal of human excellence. Rawls wants his conception of rationality to be morally-neutral. Assuming that such a neutral conception of rationality is indeed possible to attain, it is not at all evident how it could be used effectively to show that a person who does not desire a well-ordered society, or one who does not act in accordance with the Aristotelian Principle, is necessarily irrational.

Another important consideration why one cannot insist that a person must desire a well-ordered society in order to be rational, is that Rawls' well-ordered society is geared towards satisfying the "wants" and "interests" of its members as distinct from their "needs"⁽³²⁾. If his well-ordered society (or social union) aims primarily at satisfying the needs of its members, then, apart from the fact that all the members would have some basic needs in common -- a factor which might encourage a social union -- each member would be obliged to desire a well-ordered society as a means of satisfying his needs.

32 It is not a mere coincidence that Rawls is rather silent on the notion of needs. If he speaks of needs (especially basic "needs") then he might be committed to the idea of a common purpose or a common end that can by itself, forge a union both in his original position as well as his subsequent well-ordered society. Individuals who have the same needs can easily be persuaded to work together in order to attain those needs. But Rawls wants to make weaker assumptions at the original position in order to enhance the justificatory power of his contract formula. Hence we are to presume that the men of his original position have individual interests and wants.

When we talk of "needs", we are referring to those essential or indispensable things that a person must have in order to fulfill himself as a person. For instance, such things as food, shelter and protection could be defined as human needs.

Whereas a person may have an obligation (at least from the point of view of prudence) to satisfy and safeguard his needs, there is no obligation on anyone to satisfy his interests. And, if Rawls' well-ordered society is only a forum for individuals to satisfy their wants, then one may still be rational even if one does not desire it.

It would seem from these considerations that Rawls does not have an adequate conception of a social union, let alone has he provided any convincing reasons why individuals ought to desire such a union if they are rational. We now turn to the second question raised earlier: why does the desire for a social union necessarily commit an individual to affirm his sense of justice?

It might be necessary to revise this question in view of some of the issues raised in the discussion of the first question. In fact, if a social union is understood in the strict sense as depicted earlier, that is, as a society of altruists who are concerned to work together to realize their common ends and to participate in the common excellences of their social nature, then we should not anticipate any reason why a member of such a union would not desire to affirm his sense of justice.

We might even suggest that in such a union, love would take precedence over the sense of justice as a regulative sentiment. But since this type of social union is not feasible under the conditions of the original position, is there yet a way of showing that anyone who desires a well-ordered society (Rawls' version) must commit himself personally to the sentiment of justice? In other words, can a person have good reasons for not wanting to affirm his sense of justice even if he does desire a well-ordered society? Perhaps the case of the free-rider egoist which was discussed in the last chapter may again be relevant to this question. Even though all the members of Rawls' social union might see and appreciate the collective good to be derived from a social union governed by a public conception of justice, a "rational" egoist may participate in the enjoyment of such good without actually affirming his sense of justice. To enjoy his share of the benefits and to be considered rational, it is sufficient that other members of the union see him as a just man.

However, while the case of the free-rider does pose a serious challenge to Rawls' theory, it can hardly be considered an appropriate ground for rejecting his argument of congruence. Rawls' initial constraint rules out the case of the egoist. He does not consider it part of his task to show how an egoist might, or should see a social union.

Nonetheless, Rawls' argument is still vulnerable. The members of his social union being rational "self-interested" individuals, the attachments they would have for their institutional arrangements can only be contingent upon the effectiveness of the latter in the realization of their interests and abilities. But it is an empirical fact that the institutional arrangements in any society cannot serve everyone's interests or abilities in the same way. Some members of the society have rather complex abilities while others have less complex ones. The ones with more complex interests would obviously make more demands on the common resources of the union in terms of their satisfaction. Now why should those persons whose interests are less complex subscribe equally to the development of those persons with more complex interests and abilities, especially when the former do not envisage any gains from the latter? Perhaps a concrete example would better illustrate this point. Let us suppose that some individuals in the union have marked abilities in some of the highly skilled professions, say music. To develop such talents, the union would have to levy heavy taxes in order to generate funds for the required facilities. Now why would those people who do not have any interest whatsoever in the outcome of these talents help to pay for their cultivation? Why would someone, whose only interest is to pick clover leaves feel obliged to pay heavy taxes that would be used for the construction of a music theatre? How does a music

theatre or the cultivation of some singer's talent become his own good? Of course, Rawls does suggest that when men are secure in the enjoyment of the exercise of their own powers, they are disposed to appreciate the perfections of others. But this assertion is not so obviously true. In fact, its reverse would seem more plausible. Often the more proficient or skilled a person is, the more withdrawn he tends to be, and the more he tends to judge everyone else by the degree of his own excellence.

Perhaps it should be pointed out that the constraint of the veil of ignorance at Rawls' original position cannot block this objection. For instance, someone might want to argue that since nobody knows what talents he might have in real life, an agreement for the public support of all possible talents could be rational at the original position. And by appealing to the constraint of strict compliance, everyone would be obliged to subscribe to the joint effort for the development of all the possible talents that might emerge in the society, after the veil of ignorance has been removed. Even if this argument stands, it is not yet evident why helping others develop their abilities under these circumstances is necessarily a good for a disinterested individual. However, as seen earlier, the constraint of the veil of ignorance is no longer operative in the argument of congruence. One of the main tasks of the argument is to defend the rationality of the thin theory without

appealing to the constraint of ignorance. The members of the social union know what their particular interests and abilities are. They are also aware of the particular abilities and interests of others in the union. Now the question is, with the rationality of the thin theory, would it always be to everyone's greater good to support the just arrangements in the social union?

It does not seem that this question could be answered in the affirmative if we are to rely on the individualistic assumptions of Rawls' contract doctrine. In view of the fact that interests vary from person to person and some interests make greater demands on the public resources than others, some people would always be in a position to make some personal sacrifice in the union. It might even be necessary for some to sacrifice some of their interests voluntarily to enable others to realize their own interests. But such personal sacrifices are only possible if the men of the union are endowed with feelings of sympathy and altruism. Rawls' men do not have sympathetic feelings at the original position, and neither has he been able to show in any convincing way why they would necessarily have these feelings in a social union which derives from the conditions of the original position. Until an adequate basis has been provided for such feelings, Rawls' second argument of congruence remains rather inconclusive.

The third and final reason which the thin theory allows for maintaining one's sense of justice is connected with what Rawls refers to as the Kantian interpretation of justice as fairness. Rawls tries to argue here that "acting justly is something we want to do as free and equal rational beings." He bases this argument on his interpretation of Kant's conception of autonomy. In accordance with this conception, we are to think of moral principles as a "legislation for a Kingdom of ends." Once we do this, we will come to see "that these principles must not only be acceptable to all but public as well." As a result they are to be agreed to "under conditions that characterize men as free and equal rational beings."

Rawls' description of his own original position attempts to capture this Kantian conception of free and equal rational beings who legislate moral principles for a Kingdom of ends. He maintains that:

... (the) desire to act justly and the desire to express our nature as free moral persons turn out to specify what is practically speaking the same desire ... They are both dispositions to act from precisely the same principles: namely, those that would be chosen in the original position. (37)

33 Ibid, p.572.

34 Ibid, p.252.

35 Ibid.

36 Ibid.

37 Ibid, p.572.

The analogy in this argument raises some interest. If the conditions of the original position can adequately describe the Kantian conception of a Kingdom of ends (a Kingdom of free and equal rational moral persons), then we might say that the desire to act justly according to the principles of justice as fairness can indeed approximate the Kantian notion of the autonomous will by rational moral agents to express their nature as free and equal moral persons. This would in turn answer some of the objections raised earlier concerning the link that Rawls wants to maintain between his theory and that of Kant. However, Rawls' analogy does not have a sound basis. His original position cannot yield the notion of a Kantian Kingdom of ends. His rational self-interested contractors are not just capable of making an autonomous choice of moral principles in a Kantian sense. They cannot even understand what moral principles are. If they are capable of making any choice at all their choice can only be heteronomous.

Thus, it would seem from the remarks made so far that the main reasons which the thin theory of the good allows for maintaining a person's sense of justice are by themselves rather inconclusive. And, with such vulnerable premises it is difficult to see how the congruence argument can succeed. However, before coming to a final conclusion on this matter, it is still necessary to see how Rawls himself considers the decisiveness of these reasons.

In considering the adequacy of the reasons allowed by the thin theory, Rawls adopts the strategy of "a balance of motives which in many ways is similar to a balance of first principles."⁽³⁸⁾ What we do here is to compare one balance of reasons with another and "if the first balance favours a certain choice,⁽³⁹⁾ then the second does also." In this exercise, the content of the principles of justice itself becomes a crucial factor. For, as Rawls says, in order to determine whether "it is for a person's good that he has a regulative sense of justice would depend upon what justice requires of him."⁽⁴⁰⁾ The standard by which the concepts of the right and the good are specified would determine their congruence. From this standpoint, Rawls claims superiority of his contract theory over utilitarianism because of the strictness of the latter.

As Sidgwick notes, Utilitarianism is more strict than common sense in demanding the sacrifice of the agent's private interests when this is necessary for the greater happiness of all. It is also more exacting than the contract theory, for while beneficent acts going beyond our natural duties are good actions and evoke our esteem, they are not required as a matter of right. Utilitarianism may seem to be a more exalted ideal, but the other side of it is that it may authorize the lesser welfare and liberty of some for the sake of a greater happiness of others who may already be more fortunate. A rational person, in framing his plan, would hesitate to give precedence to such stringent a principle. It is likely both to

38 Ibid.

39 Ibid.

40 Ibid.

exceed his capacity for sympathy and to be hazardous to his freedom. Thus, however improbable the congruence of the right and the good in justice as fairness, it is surely more probable than on the utilitarian view. The conditional balance of reasons favours the contract doctrine. (41)

It is not very clear here whether Rawls is merely trying to provide a rationale for those main reasons which the thin theory of the good allows for maintaining one's sense of justice, or whether he is trying to set up a new and independent ground for the congruence of the right and the good. The link between this passage and those chief reasons is not exactly evident. However, whatever Rawls' actual intentions are, this passage does make several interesting claims. In fact, if some of these could be adequately sustained, we might be encouraged to take a second look at the entire congruence argument.

The central issue in this argument is Rawls' distinction between natural duties and beneficent or supererogatory acts. If this distinction could be validly sustained, and if it is indeed the case that utilitarianism makes the latter obligatory, then Rawls would have evidently made a strong case against utilitarianism and the conditional balance of reasons would seem to favour his contract doctrine. To adjudicate this matter effectively, it is necessary first to recall how Rawls characterizes natural duties and supererogatory acts.

41 Ibid, pp.572-3.

The following are examples of natural duties: the duty of helping another when he is in need or jeopardy, provided that one can do so without excessive risk or loss to oneself; the duty not to harm or injure another; and the duty not to cause unnecessary suffering: The first of these, the duty of mutual aid, is the positive duty in that it is a duty to do good for another, whereas the last two duties are negative in that they require us not to do something that is bad. (42)

In contrast with obligations, natural duties apply to every individual unconditionally.

... they apply to us without regard to our voluntary acts. Moreover, they have no necessary connection with institutions or social practices; their content is not, in general, defined by the rules of these arrangements. Thus we have a natural duty not to be cruel, and a duty to help another, whether or not we have committed ourselves to these actions. It is no defence or excuse to say that we have made no promise not to be cruel or vindictive, or to come to another's aid. A further feature of natural duties ... is that they hold between persons irrespective of their institutional relationships; they obtain between all as equal moral persons. In this sense the natural duties are owed not only to definite individuals, say to those co-operating together in a particular social arrangement, but to persons generally. (43)

Rawls says very little about supererogatory acts. He broadly refers to these as "permission":

These are acts of benevolence and mercy, of heroism and self-sacrifice. It is good to do these actions but it is not one's duty or obligation. Supererogatory acts are not required, though normally they would be were it not for the loss or risk involved for the agent himself. A person who does a supererogatory act does not invoke the exemption which the natural duties allow. For while we have a natural

42 Ibid, p.114.

43 Ibid, pp.114-5.

duty to bring about a greater good, say, if we can do so relatively easily, we are released from this duty when the cost to ourselves is considerable. Supererogatory acts raise questions of first importance for ethical theory. For example, it seems off hand that the classical utilitarian view cannot account for them. It would appear that we are bound to perform actions which bring about a greater good for others whatever the cost to ourselves, provided that the sum of advantages altogether exceeds that of other acts open to us. (44)

To begin with, each of these three passages does contain some important truth. For instance, it seems quite plausible, as Rawls suggests in the first passage, that even though mutual aid might be properly considered a natural duty for all rational persons, (as human beings we have needs which require the co-operation of others to fulfill), no person can be asked to help another at his own peril. For instance, it is hardly considered a failure in duty or love or rationality should a parent refuse to donate any of his vital organs to save the life of his child. It seems also reasonable to a point, to speak of natural duties as unconditional. It might be something of a conceptual muddle should one describe a particular act as a natural duty and yet try to define its content or force of obligation in terms of common consent or particular social practices. If anything is a natural duty, then its source and content cannot derive from social conventions or mores. Finally, it does seem that there is some basis for a distinction

44 Ibid, p.117. (emphasis mine).

between natural duties and beneficent acts as Rawls suggests in the third passage. For example, when a by-stander plunges into an inferno in an attempt to rescue a trapped victim, we could hardly be mistaken in describing such an act as heroic. But while we do recognize such an act as being good and laudable, no one expects it to be performed by everyone from the standpoint of duty (natural). On the other hand, if some parents fail to provide for the needs of their off-spring, they are said to be failing in their duty to the child. Providing for the needs of offsprings can hardly be disputed as an unconditional natural duty of parents.

However, while considerations such as the foregoing tend to support Rawls' distinction between natural duties and beneficent acts, there are still some fundamental difficulties which he must come into grips with. The first, and perhaps the most crucial difficulty is the basis of the principles of natural duties and beneficence. What are the principles that govern these types of acts? Rawls does not attempt to answer this question because his theory is ill equipped to deal with it. Admittedly, it might be rather difficult to subsume what he classifies as positive and negative natural duties under one principle. But then if one principle cannot do the job, why does he not appeal to two? It seems much preferable to explain these duties with two or more principles than to leave

them unaccounted for. But Rawls is unable to account for these principles because some of the things he describes as "positive natural duties" are grossly incompatible with the spirit of his contract doctrine. Indeed, the men of the original position may not have the tendency to injure one another or to harm the innocent. Apart from the fact that they are neither jealous nor envious of one another, refraining from injuring or harming the innocent could well be a very effective means of safeguarding themselves from injuries from others. They may also uphold the principles of justice, (as discussed earlier) to the extent that they see the principles as effective means of adjudicating their competing claims. The duty of mutual respect does not pose any problem either. Even if the contractors are said to be self-interested individuals, they may yet see that everyone can better enhance his own self-respect by respecting others. Now, even if one is prepared to ignore the fact that the duties prescribed here are purely prudential (they are dictated by the rational self-interest of the contractors), Rawls is yet to explain why his rational mutually disinterested individuals would want to endorse the principle of mutual aid, when their individual self-interest is not at stake. Rawls seems rather confused on this point. First he claims that the principles of natural duty are derived from a contractarian point of view: They are among the principles that would be acknowledged in the original position. Hence, like the two principles of justice as fairness, we can consider them established insofar

as we are able to show that the men of the original position would be unanimous in their choice of these principles. But then Rawls also insists in the same breadth that these principles "do not presuppose an act of consent, express or tacit, or indeed any voluntary act, in order to apply."⁽⁴⁵⁾ How do we reconcile these two claims? In what way are these principles yet derivable from the contractarian point of view, when the factor of consent is totally irrelevant in their application? Of course, Rawls does not want us to over-emphasize the contractual nature of justice as fairness. As he says, "once the full set of principles, a complete conception of right, is on hand, we can simply forget about the conception of the original position and apply these principles as we would any others."⁽⁴⁶⁾ This constraint is not unreasonable. Afterwards, the original position is only a hypothetical device for deriving these principles. While accepting this fact, it is not yet clear how the principles of natural duty can be shown to be consistent with justice as fairness. The former, as specified by Rawls himself are "unconditional principles" of practical reasoning. They are by their very essence "other-regarding." But for something to be considered a good by Rawls' men, it must at least have the capacity to satisfy "individual" self-interests

45 Ibid.

46 Ibid.

or further "individual" rational life plans. In a sense, some of the things listed as natural duties, especially the negative natural duties, may yet be considered good by Rawls' men to the extent that they create favourable conditions for every individual to pursue his rational life plan. For instance, by mutually acknowledging the duty not to harm one another, each member of the original position would be able to engage in his own rational life plan without fear of being harmed by someone else. While such an acknowledgment appears quite reasonable, can it be said that the contractors are motivated by a sense of 'natural' duty (a duty that binds unconditionally)? If someone (an egoist through and through) were to have some mysterious powers that makes him totally invulnerable, would there still be a need for him, within the framework of the original position, to endorse the natural duty of not harming the innocent?

Again, we are confronted once more with one of the fundamental puzzles in Rawls' theory of justice, that is, his attempt to derive the principles of a normative theory of justice and morality from conditions that are similar to the features of a private society. Rawls himself characterizes the chief features of the latter as follows :

... first that the persons comprising it, whether they are human individuals or associations, have their own private ends which are either competitive or independent, but not in any case complementary. And second, institutions are not thought to have any value in themselves, the activity of engaging in them not being counted as good, but if anything as a burden. Thus, each person assesses social arrangements solely as a means to his private aims. No one takes account of

the good of others, or of what they possess: rather, everyone prefers the most efficient scheme that gives him the largest share of assets. (47)

It is not impossible for some code of conduct such as the principles of "justice as fairness" to emerge from these thoroughly individualistic conditions. It was mentioned earlier that individuals under these conditions may yet find it necessary to engage in some form of rational co-operation in order to develop their particular interests and abilities. But the very existence of rational co-operation would presuppose some commitment among these individuals. There can hardly be a meaningful rational co-operation, if conditions of mutual respect do not exist. Persons who would not hesitate to hurt one another cannot rationally co-operate on any venture. However, whatever commitment one may see among Rawls' contractors, one can only view it from a pragmatic or utilitarian framework. At best it can only be seen as a prudential value (a means of attaining individual interests and plans), and hence, it can only apply conditionally. Such a commitment does not attain the status of the Kantian (categorical) moral imperative.

It would seem from the objections raised so far that the congruence argument does not succeed. The balance of reasons that Rawls tries to maintain in favour of his contract doctrine does not hold. His doctrine does not avoid some of the thorny difficulties that plague utilitarianism. Neither Rawls nor his utilitarian counterparts can claim to have rendered an adequate account of "unconditional principles of mora-

lity". The individualistic framework from which they both operate cannot yield such principles. Rawls is correct in noting that utilitarianism naturally presupposes sympathy and fellow-feeling. But, in actual practice, it operates on the basis of individualism. However, while recognizing this pitfall, Rawls himself has not avoided it. His contract doctrine also operates on the basis of individualism. Thus, even if his "rational contractors" would hesitate to give precedence to so stringent a principle as utilitarianism, this is not because they are 'rational' agents or that utilitarianism is irrationally stringent. Rather, it is because they are just not disposed to see the good of the other person as a sufficient motive for action. Yet taking the good of others into account is crucial to the very notion of morality. One of the characteristics of a moral person is the genuine concern for the good of others. Hence one may equate a "moral commitment" with a "social commitment"; and, a "morally good person" with a "selfless" person. The willingness to submit oneself to the pursuit of the good either on account of the 'self' or the 'other' is implicit in the notion of a moral person. One needs such a framework to be able to speak coherently of "unconditional natural duties". To be able to recognize the principles of natural duties, one must recognize at least one moral duty: the duty to have a genuine concern for the good of others. Since the latter is ruled out in the original position, the contractors cannot recognize the principles of natural duty. This claim will be carefully expounded in the next chapter. Meanwhile, a brief summary and conclusion may help put the various aspects of Rawls' theory of the good in proper perspective.

SUMMARY AND CONCLUSION

Is there a single property or group of properties by virtue of which all intrinsic or inherent goods are good? This question has raised considerable interest among philosophers all through the ages. While Rawls does not address himself explicitly to it, his theory of "goodness as rationality" may yet be construed in a very important way as one form of answer to the question. In "goodness as rationality", Rawls isolates rationality as the most fundamental concept in the notion of goodness (perhaps as one of the principal good-making properties). Thus for any given object X, X is said to be a good X if and only if X possesses to a higher degree than the average or standard X, those features that it is rational to want in X. And from the moral standpoint, a person is considered a morally good person if and only if he has to a higher degree than the average person, the features of moral character that it is rational for members of a well-ordered society to want in their associates.

Rawls develops his theory of the good in two parts. The first part which he calls the thin theory of the good, focuses on the minimal conception of the good that must be on hand in order to generate the principles of justice. Although Rawls associates his account with the deontological tradition -- maintaining the priority of the concept of right over the good -- he is aware that in order to generate the

principles of right, an appeal must be made to some notion of goodness. We would need to make some assumption about the motives of the parties in the original position. But to ensure that these assumptions do not jeopardize the priority of the right over the good, the theory of the good used in arguing for the principles of justice must be restricted to the bare essentials. Rawls refers to this account of the good as the thin theory. Its role is to secure the premises about primary goods required to arrive at the principles of justice. Once this theory is established and the assumptions about primary goods accounted for, we may then proceed to use the principles of justice in the further development of the second part of the theory of the good. Rawls refers to the latter as the full theory of the good.

The thin theory assumes that there are certain types of goods (primary goods) that all rational men must desire, if they are to desire anything at all. Whatever a person's individual life plan is the primary goods are the things that he must have in order to realize his life plan. The primary goods do not necessarily implicate a relationship to particular ends or goals; rather, they are just the ordinary indispensable means to achieve any substantive end whatever. Thus, for a person to be rational and to want anything at all he must want the primary goods.

To establish the truth of this claim, it is necessary to examine the various items on Rawls' index of primary goods. Rawls distinguishes between two basic types of primary goods: the primary social goods and the primary natural goods. The primary social goods include such things as rights and liberties, opportunities and powers, income and wealth, and the good of self-respect. The primary natural goods, on the other hand, are such goods as health, vigour, intelligence and imagination. Rawls does not speculate very much on the primary natural goods; rather, he focuses mainly on the primary social goods.

Thus the first important question that must be asked is whether it can be asserted categorically that any rational person must want any, or all of the items listed as primary social goods in order to be rational? Is it actually the case that these items are indispensable no matter what sort of an ideal form of life a rational individual chooses to pursue? It does not appear that this question can be answered in a direct affirmative. From the discussion in Chapter Four it is not evident that there are many specific things that we can quite uncontroversially claim that any rational man must want if he wants anything at all. Even within the western liberal society where most, if not all, the items listed by Rawls tend to be very pervasive, there are several individuals or groups of individuals who do not desire any of these things (at least not in the manner conceived by Rawls). But we

cannot refer to all such individuals as being plainly irrational simply because they do not desire these goods.

One important weakness in Rawls' account is his failure to make precise the context in which the items in his index function as primary goods. Perhaps within a given context, we may be able to make a reasonably strong case for some of the things he has in his index. For instance, if viewed from the standpoint of what we commonly refer to as "fundamental human needs", such things as income and wealth (defined as the legal command over the material means in general necessary to realize people's needs and interests) may indeed be the things that rational persons ought to desire. If certain things can indeed be specified as primary or fundamental needs of man, then a person can hardly be called rational if he does not desire those things that truly satisfy such needs.

However, the notion of need plays very little role in Rawls' actual account. In fact, Rawls might not consider the notion to fit very well with his contractarianism. He might even see it as a sort of limitation on the freedom of choice of his contractors. In order to create what he takes to be an ideal contract situation, he assumes, among other things, that there is a divergence of fundamental interests and ends, and a variety of opposing and incompatible basic beliefs among his contractors. Perhaps a good case may yet be made for

these constraints. There is a general tendency to laud an agreement that is reached among individuals who have serious disagreement (individuals who disagree on matters of principle). Henry Kissinger earned great respect on the international scene, and was duly honoured with a Nobel Peace Prize, because of his remarkable ability in negotiating an agreement in the Vietnam War despite the fundamental ideological differences among the parties involved in the dispute. Yet, apart from the factor of respect or elegance, it is not so evident whether conditions of divergence of fundamental interests and ends are intrinsic to the nature of a contractual agreement. Why is a contract any less a contract should the parties be presumed to have compatible aims, identical needs and beliefs? When two people come together to enact a marriage contract do we not often presume that they have compatible aims, needs and beliefs?

Of course, there is the question of what Rawls labels as the circumstances of Justice. Following Hume, Rawls insists that the circumstances of justice obtain whenever mutually disinterested persons put forward conflicting claims to the division of social advantages under conditions of moderate scarcity. While this assertion seems defensible, at least in part -- the idea of justice would hardly be relevant in a society of angels where everyone has natural commitment to the well-being of the other, or in a society where everyone can have his full in regard to the natural resources -- it is not so evident how such conditions should

necessarily rule out the notion of common needs. Why would free and equal persons who have the same basic needs not make claims and counter claims on one another? The mere fact that individuals are aware of their common needs and goals does not suggest a willingness for rational co-operation or the desire to treat each other justly.

Thus it would seem that the notion of need is not necessarily incompatible with the idea of contract or the circumstances of justice. Rawls' reluctance in deploying this concept for a more important role in his account may have stemmed from his desire for elegance. He wants to make only those assumptions that rational persons would easily agree upon. His theory would thus manifest a certain degree of elegance, should he be able to generate his principles of justice from such uncontroversial assumptions. Admittedly the notion of need does have its own peculiar difficulties. It is often asked, for example, whether we do in fact have any clear conception of what the primary or fundamental needs of man are. Also such a question as the relation between "X is good" and "X satisfies a need" must not be ignored. However, some demythologized variant of this concept may still provide a good basis for Rawls' conception of primary goods. Without placing these goods in some specific context, it is difficult if not impossible to determine whether they are things that all rational persons must want if they are rational.

Another important feature in the thin theory of the good is the notion of rationality expounded by Rawls in it. This is the centre-piece of the entire theory.

According to Rawls' account, the primary goods are said to be those things which any rational person would desire to have in order to fulfill his life plans. Apart from finding out whether the primary goods actually possess such characteristics by virtue of which every rational person ought to find them desirable, there is a more fundamental question of determining who rational persons are, or what makes one a rational person. Rawls does not attempt to propose a substantive definition for the notion of rational person. This might be a topic for meta-ethics, perhaps. While admitting the relevance of meta-ethics in ethical theory, his aim in A Theory of Justice is to construct a substantive (normative) theory of morality. However, Rawls does provide some useful insight to the notion of a rational person in his thin theory of the good. In this account, he tries to show that even though different people may want different things, there are certain standardized means to want fulfillment. He attributes to each person a "rational plan" which defines his main expectations in life. He then goes on to define a man's "good" as "the successful execution of a rational plan of life". Within this framework,

a rational person may be described as one who pursues that plan of life which enables him to realize his conception of the good. And, since the primary goods are said to be necessary means for the realization of any rational life plan, a person ought to desire them in order to be rational.

The notion of rationality emerging from the thin theory is the same as the one which is standardly employed in economic theory. It identifies rationality with utility-maximization. This conception of rationality raises several important questions: how does it fit in with our ordinary conception of rationality? For instance, does it adequately portray who a rational person is in the various contexts in which we apply the term 'rational' to persons? Does it provide an objective and acceptable standard for distinguishing between a "rational" and an "irrational" life plan? And more importantly, can such an instrumental, morally neutral conception of rationality ever lead to a non-consequentialist or intrinsic morality? These questions are the focus of our discussion in Chapter Five.

The notion of rationality in the thin theory does indeed capture a very important sense of how the term is applied to persons. If a person actually desires some object or an end, he can only be said to be rational if he acts on those plans that ensure the realization of it. For example, an athlete who wants to excel in his sport cannot afford to be indifferent to training. By taking his training seriously, he is said

to take the most effective means toward the realization of his goal. From this standpoint he may be referred to as a rational athlete.

However, while the conception of rationality in the thin theory seems quite familiar, it is not plainly evident whether it can provide adequate basis for identifying a rational person in every instance, especially from the standpoint of a rational moral agent. It is basically a rationality of a means-to-ends; a conception of rationality as distinct from the concept of rationality. But the notion of a rational person extends beyond such a narrow conception. Sometimes, to determine whether or not a person is rational, not only is it relevant to examine how he chooses or executes his plan of life, rather, what he takes to be a plan of life is itself called into question. Both the means he adopts for the realization of his goal, and the goal itself must be seen to be rational.

Rawls tries to deal with the latter by appealing to the principles of rational choice and deliberative rationality. Before a plan of life can be said to be rational, its choice must be motivated by the principles of rational choice and by deliberative rationality. However Rawls seems to doubt the adequacy of his principles of rational choice (the counting principles as he calls them) as a basis for ordering plans in accordance with their rationality. As he points out, the means for the realization of any given plan may not be neutral, inclusive

plans may not even exist, and the objectives of an agent may not be sufficiently similar. And more importantly, the principles merely provide a strategy for counting existing desires and wants in an attempt to see how many of these could be realized. They do not provide any insight in regard to the quality of these desires. Thus we must focus more on the notion of "deliberative rationality".

By appealing to this notion Rawls wants to claim that the choice of plan made by an individual under "full information" is necessarily rational, and hence objectively good. Such a plan is the objective good for a person. Again, certain things could be said in favour of this notion. The factor of full information can hardly be ignored in the assessment of the rationality of a choice. However, the notion cannot offer all that Rawls wants to derive from it. We sometimes argue -- and plausibly too -- that even though an agent wants a certain object, knowing all the facts about the object, yet "it is not the best or the most rational thing for him to want. Besides, it is not a conceptual anomaly to conceive of a thoroughly rational plan of life which is totally evil or at least not good. Thus, while a person needs to have full information in order that his choice might be considered rational at all, this condition alone cannot determine the rationality of a choice. A choice made under full information may yet be irrational and evil. For instance, Hitler may have had people

who gave him full and accurate information which he could have listened to and understood but yet he believed that Jews were evil swines to be killed to preserve the German culture. Must we now say he is rational? Thus in regard to one of the questions raised earlier, it would seem that the conception of rationality in the thin theory of the good does not adequately specify the notion of a rational person. And, until we have a firm assurance of what, or who a rational person is, it is difficult to assert categorically that he would always desire the primary goods.

Besides its inadequacy in specifying the notion of a rational person, Rawls' received view of rationality also raises concern in regard to the problem of generating the principles of justice. It is not immediately evident how such an instrumental conception of rationality can yield an intrinsic morality (a morality that can generate its own support). The men of Rawls' original position might see the value of rational co-operation. Along with this, they might desire that some principles of conduct be publicly acknowledged to regulate their competing claims. But their sense of morality cannot transcend the morality of economic men: here everyone tries to be moral in order to attain a greater well-being for himself. There is no doubt that such a morality would not be acceptable to Rawls. As a deontologist, he would like to argue that the concept of right is not only prior to that of the good, but that it has

an intrinsic value. A rational person would, after considering all the relevant facts, come to appreciate the principles of justice and morality for its own sake. Perhaps something can be said for this claim. However, it cannot be established on the received view of rationality that Rawls advances in the thin theory of the good.

The weakness of the thin theory of the good, especially the received view of rationality contained therein, becomes more apparent when we take a look at the full theory of the good.

Having defined a person's good as the successful execution of a rational plan of life, and his lesser goods as parts thereof, Rawls then goes on to apply this conception of goodness in what he calls the full theory of the good. The full theory attempts to deal with other subjects that have an important place in moral philosophy. In particular, it seeks to provide a basis for defining the moral worth of persons and for defining beneficent and supererogatory acts. Also, Rawls hopes to be able to determine on the basis of the full theory whether or not being a good person is always a good thing for a person.

In developing the thin theory into the full theory, Rawls postulates the fundamental moral virtues as the properties or features of moral character that members of a well-ordered society would rationally desire to find in one another. Therefore, should any member of a well-ordered society possess these

virtues to a higher degree than the average member, the same shall be deemed a good person. The key ideas in this definition are the notions of a well-ordered society and the fundamental moral virtues. How does Rawls account for them?

In view of his claims to objectivity and universality (morality ought to generate its own support) Rawls would want to say that both a well-ordered society and the fundamental moral virtues belong to the class of intrinsic goods. They are things to be desired by rational persons because of their intrinsic worth (goods to be sought for their own sake). If a person were to be rational and desire anything at all, then he should desire a well-ordered society and the fundamental moral virtues, especially the sense of justice.

Perhaps it is possible to show that a well-ordered society as well as the fundamental moral virtues are intrinsic goods that every rational person ought (in a non-moral sense) to seek. But Rawls has not shown us why this is so; nor does he have an appropriate basis to support such a claim. From the standpoint of the motivational assumptions of the original position, and especially from the received view of rationality of the thin theory, the most that Rawls could claim is that rational persons who are concerned to further their individual conceptions of the good will desire a well-ordered society as the most effective means of attaining their ends. Also, they will find it collectively rational that each member of a well-ordered

society should possess the fundamental moral virtues -- especially that each one should affirm his sense of justice -- as a means of effectively regulating their competing claims. But neither the notions of a well-ordered society nor that of the moral virtues operate as intrinsic goods in these claims. Rather, they operate as utilitarian or instrumental values. As such, it cannot be categorically asserted that a person will be faulting his rationality, if he, for some reasons, does not desire them. The rationale for these goods is contingent upon what an agent has taken to be his rational life plan. Apart from the theoretical possibility that an agent's human good (the successful execution of his rational life plan) may not require either a well-ordered society or moral virtues, there is a very important distinction that Rawls does not seem to take into account. That is, what is collectively rational for an individual to do is one thing; and, what is individually rational for a particular individual in a reasonably well-ordered society to do is another. It might be collectively rational for individuals to desire a well-ordered society together with all the moral sentiments which regulate it. But this does not necessarily imply that every individual in a well-ordered society will find it individually rational to uphold the moral sentiments that regulate the society. If the overriding factor is the successful execution of a life plan, then any means at all that would most effectively attain this good must

be preferred. Thus, depending on the prevailing circumstances, an individual may desire to be a morally good person in order to attain his own good, while another person may find it sufficient to be a man of good morals. Since each of them effectively attains his own good, we do not have a basis for even distinguishing one from the other, let alone preferring one to the other.

But Rawls wants to argue (as he does in the congruence argument) that at least under the circumstances of a well-ordered society, justice and goodness are congruent. That is, in a well-ordered society and perhaps more generally as well, having a sense of justice is a good, and that, since this is so, the having of moral sentiments is desirable from the point of view of rational persons who have them when they assess their situation independently from the constraints of justice. In which case, rational persons will yet come to appreciate that there are good and sufficient reasons, apart from the constraints of morality, for choosing to live by these constraints and wishing to have moral sentiments, as a well-functioning part of one's psychological make up, such that a person is not only a man of good morals but that one is as well a morally good man. Perhaps this might be so, although Rawls, as mentioned above, has not shown why it is so. Assuming however that this claim is true, does it necessarily prove the case that to be rational, or at least to be fully

rational, one must actually have these moral attitudes? Again we need to recall Rawls' received view of rationality.

Rawls does not rest his claim of congruence entirely on the principles of rational choice together with the concept of rationality. Both of these, as he argues, cannot by themselves sufficiently determine what would be a rational and indeed a good life plan for a human being. Rather we must take into account some crucial general facts about human needs and abilities. We are to rely of course, on our common sense knowledge in order to specify these crucial general facts about human needs and abilities. Thus, given his definition of good (what it is rational to want), his conception of rationality and reasonable careful reflection (simply using our common sense knowledge about human wants, needs and abilities), Rawls hopes that we will arrive at some fundamental human goods such as personal affection, friendship, meaningful work, social cooperation, the pursuit of knowledge and the fashioning and contemplation of beautiful objects.

Indeed it seems quite plausible as Rawls points out that these familiar goods are activities and ends, which have an important place in the lives of human beings. However, it is not immediately evident how achieving these ends would necessarily involve an intrinsic morality. For instance, why would a classist amoralist not equally attain them? Why can he not limit his affection and friendship to the members of his own

class? Why is personal affection or friendship necessary for social co-operation?

The fundamental weakness of Rawls' theory as it emerges from this survey is that his premises are too weak for the claims he tries to derive from them. He tries to get more out of the received view of rationality vis-à-vis morality than the conception can yield. There are several constructive insights in Rawls' theory no doubt. His principles of justice seem to square, under certain conditions, with our ordinary considered judgments about what is just, and his account of the psychology of moral development is very impressive. However, there is a need to strengthen the foundation of the theory for a better consolidation of these insights. A slight revision is desirable at the original position in order to fortify the notion of a well-ordered society. Rawls would need an enriched notion of rationality and, perhaps, certain considerations of altruism. The possibility and extent of such a revision will be examined in the next chapter.

PART III

TOWARDS A REVISION OF RAWLS' ACCOUNT

CHAPTER VIII

SOME PROPOSALS TOWARDS A RECONSTRUCTION OF RAWLS' ACCOUNT

This study has so far revealed several difficulties inherent in Rawls' theory. Perhaps the most fundamental of these, is his attempt to proceed from considerations of rational self-love (prudential concerns) to establish the motivation to act from the moral point of view. Rawls has not succeeded in doing so; nor is it certain that anyone can. A motivation of self-interest cannot yield a moral motivation since there are no non-moral reasons sufficient to require a commitment on the part of a rational person to the moral point of view. Also, from the standpoint of the moral sceptic, Rawls' account has not been able to show in any convincing way why acting from the moral point of view is always the rational thing to do (rational in a morally neutral sense), even when the requirements of morality are directly in conflict with those of rational self-interest. More importantly, his theory of Goodness as Rationality does not provide adequate criteria for discriminating between rational persons or ends; nor does it show why it is more intrinsically desirable for a person to want to be a morally good person and not just a person of good morals.

For Rawls, a morally good person is one who has an effective sense of justice and acts from principles of justice

when they relevently apply. This characterization is by itself quite tenable. It does portray a very important sense of what we mean when we refer to a person as being morally good. However, Rawls' substantive account of the development of the moral personality fails to provide adequate basis for discriminating between a person who genuinely desires to act from the motivation of justice and a person who merely desires to act according to the principles of justice just to promote a private selfish end. One can, in certain circumstances, effectively appeal to justice just to promote a selfish end. (Recall Shylock's appeal to justice in support of his demand for a pound of human flesh in Shakespeare's Merchant of Venice.)

The distinction between a morally good man and a man of good morals is very important in the dimension of moral assessments. It is not a sufficient condition for referring to someone as a morally good person just because he performs an act that can be judged as morally good. Consider a man who plunges into a river to save a drowning boy just because he wants to receive a citation of honour. Even though his action of saving a drowning boy could be judged as morally good (it is the right thing to do), he cannot be praised for it, if his motive were known. Instead, he might even be criticized. Now, since the constraint of neutrality in Rawls' thin theory does not allow us to raise questions about motives,

desires and ends (these must be taken as given both prior to the choice of principles of justice as well as the application of these principles), there is no sure way of establishing whether or not a person has a genuine or effective desire to act from the moral point of view. As such, a very important question in moral assessment, the question of praise and blame cannot be entertained within the framework of Rawls' account.

Perhaps one might ask why Rawls ought to address himself to issues such as these. To answer this question, we need only recall some of the reasons Rawls gives why it is necessary to have a full theory of the 'good'. In developing the thin theory into the full theory, Rawls wants to be able to deal with some of the issues that are central to moral philosophy. One such issue as he does acknowledge himself, is whether "being a good person is a good thing for that person"⁽¹⁾. Rawls' insights on this question can hardly be adequate if his account does not even provide an effective strategy for distinguishing a genuinely good person from an opportunist who merely plays the role of a good man.

The issues raised above are also relevant in view of the type of ethics that Rawls sets out to do. He wants to do normative ethics in a very important way; a way that is

(1) Rawls, A Theory of Justice, p.397.

different from such prescriptivist account like Hare's. Rawls is not just trying to discover what a person's moral principles are by studying what he does, or how he uses moral concepts. Rather, he wants to specify, through a process of rigorous moral geometry, what rational and objective moral principles ought to be. He wants to provide an Archimedean point for judging the whole of morality, that is, the design of social institutions and whole moralities. To be able to deliver on these claims, Rawls cannot avoid the questions raised above. However, as this study has shown so far neither his direct nor indirect attempts to deal with them have been successful through and through.

Thus, one question I want to investigate in this concluding chapter is whether there is a way to overcome at least some of the main difficulties that plague Rawls' account, while retaining at least most of the central features of the original framework of his theory. In other words, is there a way of reconstructing his account such that his principles can be made to rest on a more secure foundation without an essential modification of his main arguments?

This question does not seem quite easy to answer. There will be the possibility of disagreement about what the central features are or what an essential modification is. Indeed, if one reasons along the line of some of the critiques that have been published on Rawls' theory, it would appear

that any attempt to reconstruct it must involve a radical revision of its main features. Perhaps this is just what an adequate revision of A Theory of Justice might come to in the final analysis. However, the scope and purpose of this study do not entail this. A defence of any particular conception of morality or way of developing a moral theory is not contemplated here. Rather, my purpose is to determine to what extent Rawls' theory, in some limited and reasonable modification of it, is able to overcome some of the central problems in contemporary ethical theory. Within such a limited aim therefore, I want to suggest one important factor that must be taken into account, if a revision of Rawls' theory is to undermine the difficulties which have emerged from my examination of his account. This is: the need to go beyond the received view of rationality (the rationality of the prudent maximizer) in the construction of a normative ethical theory. To be more specific, there is a need for a conception of rationality that involves some of the commonly accepted ideals of moral personality within the framework of a theory of morality.

For obvious reasons a morally built-in conception of rationality will not be acceptable to Rawls. Apart from the controversies that such a conception of rationality might generate, the deductive strategy of his theory would be rendered ineffective should he allow substantive moral claims in the original position. For instance, if it is the case

(2)

that his conception of rationality already entails an appeal to some moral ideal, why can the sceptic not rationally (in a non-moral sense) demand why he ought to be rational in such a sense? Again, one might wish to dismiss this question on the ground that it is not Rawls' intention to deduce the constraints of moral commitment from the requirements of 'means-end' rationality. Hence he does not have any business explaining to the sceptic or egoist how adherence to the principles of justice or morality will further his purposes. Rawls' task is only to try to provide persons who already have a sense of justice with a clearer understanding of what acting justly requires of them by articulating the correct principles of justice. Some of the preliminary remarks by Rawls in his argument of congruence seems to support this reading. And, if this is in fact all that Rawls wants to accomplish, he may still be making an important contribution to ethical theory. But then, what is the point of focussing the original position on considerations of self-interest if Rawls is merely trying to explain to the just man what justice requires of him? Are the concerns of self-interest not one of the factors that often clouds a man's vision as to what justice and fairplay require of him ?

2 Recall that for Rawls, the superiority of A Theory of Justice over other contract and teleological theories lies in the fact that its premises do not make controversial claims. The assumptions in the original position are a constrained minimum.

Perhaps judgment ought to be reserved on this question until the case for the modification suggested above has been presented. However, if Rawls' account has not been able to undermine the challenge of moral scepticism despite its non-controversial conception of rationality, it does not appear that he has too much to lose should he introduce a stronger, more controversial but plausible theory of rationality. The idea of the original position seems to raise a classic dilemma. Should Rawls introduce substantive moral notions into it, it would lose its justificatory force, at least in the deductive way he wants to use it. Above all, the moral sceptic cannot be persuaded in a non-trivial manner that his good consists in endorsing and acting from the moral point of view. On the other hand, if Rawls does not introduce moral notions into the original position, he cannot arrive at the moral point of view.

Be it as it may, we shall proceed to examine the revision that is being proposed here. We shall examine the rationale for a morally built-in conception of rationality. The strategy here is to sharpen and make more precise some of the arguments presented earlier in Chapter Five. It will be argued that a neutral conception of rationality cannot yield all that Rawls hopes to get from it. More specifically, it will be argued that unless Rawls introduces a conception of rationality that allows him to challenge and criticize the goodness or rightness of individual desires and ends, the egalitarian theory of justice

he defends in justice as fairness cannot yield any spectacular social reforms. As it stands, in a society that he would regard as well-ordered (society in which his two principles of justice are satisfied) one may still find intolerable differences in the life prospects between the advantaged strata of the society and the least well off. Indeed, a Rawlsian well-ordered society may still be morally repugnant through and through.

BEYOND THE RATIONALITY OF THE PRUDENT MAXIMIZER

Although Rawls wants the description of the original position to capture the idea of a well-ordered society (a society in which individuals are to view themselves as free and equal moral persons), his assumptions must be general and as morally neutral as possible to ensure that the principles chosen by the contractors do not favour the ideals of any particular society or groups of individuals. Hence, as stated earlier in Chapter Five, the parties in the original position are said to be rational to the extent that they are capable of taking the most effective means to attain their desired ends. This economic model of rationality does not, according to Rawls, involve any controversial claims. Its demand is simply that a person should structure his ends coherently and pursue them effectively. Also, it requires that ends be adjusted where it is feasible to maximize satisfactions. The parties in the

original position see themselves as rational in this sense, and they expect that persons in real society would also be rational in this sense whatever other traits they might have.

One can easily adduce two reasons why this notion of rationality would appeal very strongly to Rawls. In the first place, since he wants to simplify the problem of moral philosophy (reducing it to the more manageable problem of rational choice) the standard conception of rationality already existing in the latter field would naturally appear handy. It would hardly make sense for him to want to convert the problem of moral philosophy to that of rational choice, and yet discard the notion of rationality contained in the latter. The second reason why Rawls would want to settle for a morally neutral conception of rationality (a thin conception of rationality as it were) can easily be located in his desire for an Archimedian point for assessing institutions and moralities. He needs a neutral conception of rationality in order to arrive at a conception of justice that can suitably be used for assessing, without prejudice, the degree in which any given society is just. If such a conception of justice is possible at all, its derivation must not be based on particular contingencies. It must not evolve from the particular arrangement of institutions; nor must it depend on particular values, interests, desires or aspirations of individuals who may have been subjected to a particular socialization process.

Rather, the conception must evolve from persons exhibiting "their freedom, their independence from the contingencies of nature and society."⁽³⁾

Rawls is aware of the influence that social institutions have on the values and desires of individual persons in society. To be able to evolve a conception of justice that would treat everyone fairly in all possible social arrangements (a neutral conception of justice as it were), he must not assume that desires or values that tend to develop under certain social systems are more valuable or worthy of satisfaction than those which tend to emerge in other systems. Accordingly, his conception of rationality which plays a major role in the derivation of his conception of justice must be seen to be neutral. Thus, Rawls must totally ignore the particular contents of the desires of the parties in his original position. No person's desires may be presumed to be more or less worthy of pursuit than those of others. Should one person

3 Rawls, op. cit., p.256.

The basis of such a conception of justice may be conceived in terms of Kant's noumenal persons. As such, its conception of rationality, as Wolff rightly observes, must derive from "timeless, purely rational agents capable of abstracting from the limitations and particularities of their empirical characters". (Robert Paul Wolff, The Autonomy of Reason, New York, Harper Torch Books, (1973), p.226.

desire to be the president of their society while another merely wants to be a clover leaf picker, both must be seen as equally rational to the extent that they take adequate and effective means towards the realization of these life plans. Neither of these two desires must be considered superior to the other, and the principles of justice to be chosen at the original position must yield a social structure that would enable each of these individuals to pursue his desires effectively. ⁽⁴⁾ Within the framework of Rawls' theory therefore, desires must be taken as given, and rationality defined strictly in terms of their coherence and likelihood of satisfaction. Not even the source of desires may be questioned.

There is no doubt that such a conception of rationality might receive some support in our society. These days, not even parents want to decide what is good for their children and force it upon them regardless of how the children feel about it. It is generally considered inappropriate in our liberal ethos for one person to say whether or not another person's desires and ends are worthy of pursuit. One of the great merits of this society, it is often claimed, is the fact that it enables everyone to do "his own thing" provided the rights of others are not violated.

4 It would appear at first that an appeal to the Aristotelian principle could provide a basis for considering the desire to become president as being superior and more rationally desirable. But this would entail defining the Aristotelian principle as partly constitutive of rationality. Rawls does not want this as it might commit him to perfectionism.

However, despite the intuitive appeal of Rawls' model of rationality, and despite its apparent consistency with the conception of justice he wants to derive, the conception yet raises a few serious questions; questions that seem to threaten the very foundation of Rawls' theory.⁽⁵⁾ As seen in Chapter Five, if by appealing to this model of rationality, Rawls wants to claim that a person can act rationally only if he seeks to maximize that which is in his self-interest, then his claim is plainly arbitrary and false. There are numerous instances in which a person's action may be considered rational even if he does not seek to maximize his self-interest.⁽⁶⁾ However, this is not the issue in contention in this section. Rather, the point to be argued here is that the model of rationality in the original position is capable of frustrating at least one of the fundamental objectives of Rawls' theory of justice.

As stated earlier, one of the principal aims of Rawls' theory of justice is to neutralize or at least minimize as much as possible the influence of social patterns and institutions on people's values, desires and aspirations. Social patterns and institutions must be effectively regulated so that individuals cannot take undue advantage of them over one another. But the thin conception of justice (a conception that

5 This conception of justice may be referred to as "a thin conception of justice" in the sense that Rawls wants it to fit all possible social arrangements.

would fit all forms of social arrangements (capitalist or otherwise) emerging from the original position is just too weak to accomplish such a task. The constraint of neutrality of the original position does not permit Rawls to raise questions as to whether or not the desires of his contractors are right and just. He must take their desires and beliefs as given, and accept their ends and life plans as rationally chosen provided these are coherent (whether these ends, beliefs and desires are internally consistent and properly ordered). Now, how can Rawls ever hope to regulate social patterns and institutions if the desires and ends which form their basis cannot themselves be challenged? If we cannot ask how desires and ends originate or how the pursuit of them is a person's good, on what basis do we proceed to regulate them through social patterns? Does Rawls not seem committed to the viewpoint that the desire or end that a person happens to have necessarily coincides and defines his good provided he pursues it effectively and he does not interfere with other people's rights and desires?

There are some passages towards the end of Rawls' book that seem to provide an affirmative answer to the last

6 See Kai Nielsen's arguments against this model of rationality in "Principles of Rationality", Philosophical Papers 69, (1972). Also in "Rationality and Egoism", Studi internazionali di filosofia, 1975.

question. In his discussion of "Happiness and Dominant ends", he admits the inability of his account to deal effectively with the question of rational life plans. How in general, he asks, "is it possible to choose among plans rationally? What procedure can an individual follow when faced with this sort of decision?"⁽⁷⁾ In replying to these questions, Rawls admits that the appeal to the principles of deliberative rationality and other principles of rational choice cannot guide us all the way through. As he says, "we eventually reach a point, though, where we just have to decide which plan we most prefer without further guidance from principle"⁽⁸⁾. What Rawls is saying here, in effect, is that somewhere along the line a person's rational choice just terminates into a purely preferential choice; that is, a choice not rationally defensible. Thus, what a rational agent rationally prefers is, in the final analysis, just what he happens to prefer. Rawls seems to suggest a move in this direction much earlier in his discussion on the good. He asserts that "there is nothing necessarily right, or morally correct, about the point of view (rationality)⁽⁹⁾ from which things are judged to be good or bad". With this

7 Ibid, p.551.

8 Ibid.

9 Ibid, p.403.

statement one begins to have a second thought if Rawls is indeed not a subjectivist at least in the sense suggested by Hare. However, this is not the issue for consideration here. The point at issue is this: if the rational choice of the parties in the original position is in the final analysis just a matter of pure preference, then one cannot readily presume that whatever principles would be agreed to by these "rational contractors" must be principles of morality or of justice. Such a claim would just be as good as saying that whatever principles these contractors happen to prefer must be regarded as principles of justice. Rawls is quite aware of this difficulty and he tries to deal with it. As he says "the concept of rationality by itself is not an adequate basis for the concept of right; and in a contract theory the latter is derived in another way"⁽¹⁰⁾. But if Rawls is so sure about this, one wonders why he insists so vigorously and persistently that the rationality of the contractors together with the veil of ignorance adequately ensures that their unanimous choice will yield moral principles. If the veil of ignorance, as he assures us, also rules out substantive moral notions, is he not forced into the naturalistic position of having to generate his moral concepts

¹⁰ Ibid, p.404.

from non-moral considerations? Besides, if the rationality of the contractors ultimately terminates into pure preference (non-rational desire) why should some of them at least, not show preference for the principle of utility instead of the principles of justice as fairness? Or, as 'rational' men even, why can they not choose the principle of average utility in place of Rawls' principles?

Rawls might respond to the first question by drawing attention to the claim that the original position includes "features peculiar to moral theory". The original position is designed such "that the principles that would be chosen (in it) , whatever they turn out to be, are acceptable from a moral point of view".⁽¹¹⁾ Perhaps the easiest way to establish the truth of this claim is to examine the features described in the original position to see whether they are indeed peculiar to morality in general, and not just the conception of morality that Rawls is interested in defending. Of course, one might ask whether such an exercise is actually necessary since the original position is said to be hypothetical or imaginary. Saying it is hypothetical may just mean that its premises are arbitrarily drawn.

However, Rawls does not intend the original position to be interpreted in this way. As he says, the assumptions in it are philosophically the most favoured interpretation of the initial choice situation for the purpose of generating a

11 Ibid, p. 120.

(12)
theory of justice". He maintains that his assumptions are natural, plausible and reasonable since they are in fact widely agreed upon and commonly shared. They are based on "general facts of human nature" and there are good reasons that evidently connect them with the purpose of moral principles.

This insistence, as Aharon F. Kleinberger rightly points out, indicates a logical shift in the main burden of Rawls' argument. (13) Instead of proceeding along the standard contractarian strategy of trying to show that his particular conception of justice has the best solution for a precise problem of rational prudential choice, he now wants to move in a meta-ethical direction. He now tries to justify his peculiar interpretation of the initial contract situation as the embodiment of all reasonable constraints on moral arguments and principles. This is a significant departure from the original strategy Rawls had wanted to use in the justification of justice as fairness. Should he proceed in this way, then he would no longer be substituting judgments of rational prudence for ethical judgments. Rather like Hare, Nowell-Smith and many others, he would be chiefly concerned with the task of

12 Ibid, p.122.

13 A. Haron F. Kleinberger, "The Social Contract Strategy for the Justification of Moral Principles", Journal of Moral Education, Vol.5, No.2, pp.107-126.

discovering what the nature and purpose of morality is, and the conditions which arguments and principles have to satisfy in order to count as moral arguments and principles.⁽¹⁴⁾

But for such a meta-ethical strategy to succeed, as Kleinberger rightly observes, Rawls must be able to do two things: first he must be able to show that the features of the original position are indeed "peculiar to moral theory". In other words he must be able to show that these features are so distinctive of moral reasoning such that one can distinguish moral arguments, principles and judgments, from non-moral ones just by taking them into account. Secondly, the conditions of the original position, as mentioned earlier, must qualify as necessary conditions of morality in general. In which case, the conditions must be sufficiently neutral to fit any conception of morality. The latter requirement must be met in order to forestall the danger of circularity. Again, Rawls' strategy fails to meet both of these conditions. Among other things, his conception of rationality is not distinctive or even peculiar to moral theory. It is, as he says, the conception of rationality that is both standard and familiar in economic theory.⁽¹⁵⁾

14 Ibid, p.116.

15 See Kleinberger's detailed comments on all the features of the original position. He argues, among other things, that the conditions of mutual disinterestedness are not peculiar to moral theory. The same assumptions are made in economic theory. Also, Rawls' formal constraints may indeed be

Let us suppose, in order to carry on our discussion, that Rawls does have some important way of responding to the criticisms that have so far been raised or that his account can be modified in inessential ways to accommodate them. Thus we shall assume that even though his notion of rationality does seem to collapse eventually into pure preference, it can still provide an adequate basis for making good prudential judgments. Also, we shall assume that even though this notion of rationality is by itself not peculiar to moral theory, there are other features in the original position that are both peculiar and can sufficiently guarantee the choice of moral principles. It has been argued above that these claims cannot be defended within the framework of Rawls' theory. However, be that as it may, there is another important criticism to which Rawls would appear quite vulnerable in view of his reliance on a neutral conception of rationality. The latter is: since the constraints of moral neutrality does not permit us to raise questions concerning the rightness or goodness of individual desires, ends and motives, Rawls' well-ordered society (a society effectively regulated by the two principles of justice as fairness) may yet turn out to be a morally

necessary for something to count as moral principles. Nevertheless, these formal constraints are also shared by all first principles in any and every field. (Kleinberger, op. cit., pp.116-120.)

repugnant society through and through. And should this be the case, the notion of a well-ordered society as it derives from Rawls' original position cannot provide adequate basis for the definition of a morally good person.

Being placed in a situation to choose the principles which are to regulate the institutions of their society, it appears at first that the representative persons of the original position could well take advantage of such a unique opportunity to fashion the outlook of their society. Since these persons are presumed to know how social conditions in their society would eventually affect the desires and ends of its members, they must also know that by choosing a certain social structure they are more or less setting the standard for the kinds of desires and ends that people may have in their society. Thus, they might want to take advantage of their unique position to set the social goals to be pursued in their society: they might use the choice conditions of the original position to pattern a social structure that is consistent with their conception of the good.⁽¹⁶⁾

¹⁶ It would appear at first that the constraints of neutrality in the original position would rule out such an outcome. Since the parties do not know their individual desires, ends and life plans, they cannot proceed directly to choose a particular social structure that would further their conceptions of the good (they do not even know what their conceptions of the good are). Nonetheless, they might still be able to influence the system of values in their society. Recall that they are to have "general knowledge" about societies. With such knowledge they could know what kind of society can effectively guarantee the satisfaction of the desires and ends that would arise in it.

Such an outcome is unacceptable to Rawls. Even though the contractors seem to embody the 'ideal essence' of unadulterated human nature (they have been stripped of all the accidental qualities of human nature, in particular they are said to be rational and free from envy), to allow them the leverage of setting the goals and desires that people may pursue would involve Rawls in perfectionist ethics. Thus to avoid this outcome he stipulates that the contractors know that they, as well as other members of their society, already have in real life, desires and ends which they would want to realize. Such desires and ends could form the basis of their individual rational life plans as well as their conceptions of the good. Even while the specific content of these desires are yet unknown, the contractors know that they would not be prepared to modify or sacrifice them unless they are in conflict with the principles of justice.

Now, if it is the case that the contractors already have their desires and plans of life prior to their entry into the original position, then whatever principles they may eventually adopt and whatever social structure derives from such principles, these cannot determine what kinds of desires and ends individuals in real society should have. Hence, the merit of the choice situation in the original contract is not one by virtue of which the contractors could mould people's desires and life plans. Rather, it is an

opportunity to create, as it were, a democratic society: a society in which everyone would be free to pursue any desires or life plans he chooses insofar as these do not violate the two principles of justice as fairness. A society in which everyone would be free, subject to the constraints of justice, "to do his own thing". Thus it would seem that even though Rawls' well-ordered society is one that is effectively regulated by the principles of justice as fairness, the substantive desires and ends that are pursued in it would depend largely on the individual preferences of its members. We might expect therefore that in such a society some individuals would want to realize their most complex abilities in accordance with the Aristotelian principles, while others may just be satisfied picking clover leaf or counting blades of grass. Also some persons may want to be altruistic, while others may just be concerned with their self-interests, provided such interests do not interfere with the rights of others. The desires and beliefs of individuals must be taken as given, and ends as based on them.

Now, if the above characterization accurately reflects Rawls' notion of a well-ordered society, then his well-ordered society cannot provide an adequate point of view for the definition of a morally good person or the moral worth of persons. It is quite conceivable, as Mary Gibson rightly argues, for such a well-ordered society to be very inegalitarian,

exploitative and morally repugnant through and through, even though its social structure meets the constraints of the two principles of justice as fairness⁽¹⁷⁾. An illustration would be helpful here.

Consider a society roughly divided into two classes A and B. Let us suppose that A is the advantaged strata of the society and B the least well off. As we would ordinarily expect, the desires and, in fact, the actual life prospects of the children in this society would be largely determined by the class into which they happen to be born. For example, an educated person, if he or she is an engineer, scientist or a professor, his or her knowledge, contacts, travel and a whole host of life experiences can hardly but put his children in a far different position than the children of an unskilled labourer or a garbage collector. Thus, while the children in the least well off class would come to desire social ideals consistent with their socialization process, the children in the privileged class could more easily conceive ideals of life that truly reflect their background. Now, if desires and beliefs must simply be taken as given (no question may be asked about their source) why can an inegalitarian society

17 Mary Gibson, "Rationality", Philosophy and Public Affairs, Vol.6, No.3, (Spring 1977), pp.193-226.

such as this, not be just and well-ordered in a Rawlsian sense. To begin with, the moral neutrality of his conception of rationality precludes us from questioning the foundational beliefs of anyone in this society. We must simply consider the beliefs and desires of individuals as given. Now, Rawls' first principle speaks of equality of basic liberties and opportunities. Let us suppose that the society has a democratic constitution that effectively guarantees this: everyone is equal before the law, and eligible for any office in the land, including the highest office of president. No member of B would, of course, aspire or even desire to be president. It would be irrational for him to entertain such a desire since he knows that no one would vote for him. But in as much as the constitution formally guarantees that everyone in the society can become anything he wants (provided the rights of others are not tampered with), the requirement of the first principle of equal liberty is apparently met. It is of no consequence whether or not everyone wants to take advantage of this principle. The case for the second principle is rather obvious. We need only recall Rawls' Summary Statement about the maximin criterion. The second principle (the difference principle), he says could be characterized from a Marxist framework as follows: from each according to his abilities and to each according to his needs. Well, since the members of B only have limited ideals, their desires, wants and needs

could easily be met. They could be made as well off as possible by being given a greater index of primary goods. It is another thing, of course, whether the primary goods would do very much for them. For instance, if they were given greater "basic liberties" without the worth of greater basic liberties, it is doubtful whether such an offer would make much sense to them. But are we to conclude then that a society such as this is just and well-ordered simply because everyone's desires, ends and plans of life are realized? Or shall we say that because someone voluntarily accepts to be exploited therefore such an exploitation is morally just and acceptable? Why could the members of B not be shown that their desire to remain at the bottom of the social ladder is irrational though this is what they want. Of course, Rawls would respond here that both the parties in his original position as well as the members of his well-ordered society conceive themselves as free and equal moral persons. But he does admit that, where his two principles of justice are taken to be rational ordering principles for the guidance of social relations, it could be the case that justice, and indeed a commitment to morality, would require the acceptance as just, and as through and through morally acceptable, a not inconsiderable disparity in the total life prospects of the children of entrepreneurs and children of unskilled labourers, even when those children are equally talented, equally energetic

and the like. A just society, he claims, could under certain economic conditions tolerate such disparities. In fact, this is what Rawls' theory commits him to. It is doubtful, however, whether such a judgment would fit the considered moral judgments of most people in reflective equilibrium. On the contrary, several people would think that such a society can not be a just society, let alone a perfectly just society.

It appears, from the remarks made so far, that Rawls' theory needs a more enriched conception of rationality. The concept of rationality is very crucial within the framework of his account. As it stands, the conception he has chosen is not very effective in view of what he hopes to get from it. If no questions may be asked about the desirableness of persons' wants and desires, apart from their internal consistency, then there is no way of ensuring that the desires and wants of the persons in a well-ordered society will indeed be rational or at least acceptable from the moral point of view. Indeed there is a real possibility, as Mary Gibson suggests, for a person to be rational (in Rawls' sense) no matter how pointlessly harmful or self-destructive his desires and ends might be.⁽¹⁸⁾ Rawls' conception of rationality only requires that ends and desires be properly ordered

18 Mary Gibson, op. cit., p.200.

and effectively pursued. But if the remarks made so far are correct, it seems quite possible to question in a very important way, the rationality of certain persons even though the desires and ends they pursue are properly ordered and effectively pursued. An adequate conception of rationality (one that is indeed peculiar to a moral theory) should enable us to project certain ideals of moral personality as being worthy of pursuit, and to condemn certain ways of life as being morally repugnant. Without such a conception of rationality it is difficult, if not impossible, to regulate the desires and wants of the members of a well-ordered society to make these coincide with the desires and wants of moral persons. Unless such a theory of rationality is on hand, Rawls' conception of a well-ordered society remains very vulnerable.

Now let us assume that these observations are correct. That is, we shall assume that a plausible case has been made here for an enriched conception of rationality⁽¹⁹⁾. There are, at least, two questions that one may yet ask: (1) how does one generate such a conception of rationality; and (2) how far can the original plan of Rawls' theory accommodate such

¹⁹ It is quite possible, of course, that Rawls might still be able to make some effective response to all the objections that have been raised. In particular, he might insist that the classist society cited earlier does not satisfy the two principles of justice as fairness.

an enriched conception? The discussion for the remainder of this chapter will concentrate on these two questions. We will begin with the first.

TOWARD AN ENRICHED CONCEPTION OF RATIONALITY

The moment one begins to conceive the persons in the original position as moral persons capable of reasoning and acting from the moral point of view, the need to fortify Rawls' neutral conception of rationality becomes rather obvious. One sees the need to go beyond the rationality of self-interest (prudent maximization) to a rationality that would constrain each of the parties to recognize that, as moral agents, it is sometimes necessary and rational to act with a direct reason to promote the good of the 'other'. But how does one establish such a built-in conception of rationality? Here, a moral sceptic might concede that the moral point of view does in fact commit a moral agent to take the interest of others into account if his action is to be consistent with the constraints of morality. He might even concede that there is no reason to think that such a commitment is irrational. Now, while making this concession, the sceptic might still be eager to know why such a commitment must be seen as part of the concept of rationality.

The task in the next few pages is to see whether some reasonable response could be made to this question. It might be useful to note at the outset that, whereas it is

possible to inspect the rationality of a particular action by appealing to some criteria of rationality in a detached and objective fashion, there is no question of 'standing outside of rationality' in order to inspect or justify rationality. Hence, the most that could be done here is to try to show that the enriched conception of rationality that is being advocated for Rawls' theory does fall within a standard of human behaviour that is commonly known and thought to be rational. In other words, the task here is to show that this conception does capture at least one of the very general things that a person would do because he is rational and which we would, on reflection, be prepared to call "a principle of rationality". As a basis for our discussion, we will consider one of the fifteen principles of rationality formulated by Kai Nielsen in his paper "Principles of Rationality".⁽²⁰⁾

The statement of the tenth principle in Kai Nielsen's index reads:

Those ends, which, from a dispassionate and informed point of view, one values absolutely higher than one's other ends, are the ends which, ceteris paribus, are to be achieved. A rational agent will, ceteris paribus, seek plans of action which will satisfy those ends; and plans to satisfy his other ends will be adopted only in so far as they are compatible with the satisfaction of those ends he or she values most highly. (21)

20 Kai Nielsen, "Principles of Rationality", Philosophical Papers, Vol. III, No.2, Oct., 1974, pp.55-86.

21 Ibid, p.58.

Before taking up the question of the claims in this principle it might be necessary first to clarify two important phrases. These are "from a dispassionate and informed point of view" and "ceteris paribus". The point of incorporating the phrase "from a dispassionate and informed point of view" is to exclude certain cases in which an agent may have a set of absolutely "valued" ends and we may still want to say that he ought not to pursue those ends. Cases of alcoholism and drug addiction are good illustrations of this constraint. Even though a drug addict or an alcoholic may value the sensation or state of 'being high' above anything else, we would still consider the cravings for drugs and alcohol as intrinsically undesirable. As D.A.J. Richards remarks;

These forms of cravings are not forms of desire which the agent values absolutely higher than his other ends from a dispassionate and informed point of view, but are precisely forms of craving which undermine the agent's very capacity for rational assessment and planning, let alone control, to the frustration of his substantial interests. (22)

Much could be said in favour of Richards here. Nonetheless, it is not so evident whether his disclamation of cravings as forms of desire does not appeal to judgment. We will come to this in a moment. Meanwhile, we take it that the very introduction of

22 D.A.J. Richards, A Theory of Reasons for Action, (Oxford), 1971, p.229.

the constraint "from a dispassionate and informed point of view" implies an element of discrimination as far as the ends desired are concerned. Thus, we might say that of all possible ends that an agent might value, only those ends which he values from a dispassionate and informed point of view need be reckoned with.

The rationale of the second constraint "ceteris paribus" is fairly obvious. There are several factors to be taken into account in the assessment of a desired end apart from the fact that an agent values it. For instance, someone who values and desired intensely to attain an end which he has no chance of attaining is no more than a dreamer. Of course, one might say that it does not hurt to dream, at least in certain circumstances. Even if this is true, there is hardly anything necessarily rational about dreams.

With these preliminary remarks, we now turn to the substantive claim in the principle itself. How does it articulate the concept of rationality? And, in what way does it transcend the Humean received view of rationality? To answer the first question, it suffices to note that one important factor which is generally taken into account when individuals are said to be rational, is the way they go about trying to achieve their desired goals. If a person is known to desire a thing intensely, we would expect him, if he is rational, to take appropriate means to attain it before contemplating anything else. Such phrases as

"primary goals", "principal objectives", as against "secondary or subordinate goals", are normally invoked to indicate which ends a rational person ought to give prior consideration when formulating his plans of action. Except when there are persuasive and overriding reasons, a person can hardly be considered rational if he continues to give prior consideration to his secondary goals when formulating his principles of action. (24)

However, it does not appear that the second question can be so easily dismissed like the first one. The central issue to be considered in the latter, is whether Nielsen's tenth principle of rationality does contain anything which is not accounted for in the received view. For obvious reasons, those who defend the received view (Richard Brandt, David Richards, John Rawls, et. al.), would insist that Nielsen's principle does not add anything to the notion of rationality. David Richards, we recall, maintains quite emphatically that questions of rationality depend entirely on the means an agent selects to attain his desires. Yet, does the very constraint "from a dispassionate and informed point of view" not itself suggest that there is more to the notion of rationality than an agent merely taking the most effective means toward his most desired ends? If Brandt's thesis that the rational act is that which maximizes the agent's long-

24 This, of course, might still be considered as part of the received view.

range expectable utility, captures all that is involved in the term "rational", why would anyone, including Richards, worry at all whether an agent's most cherished desires are from a dispassionate and informed point of view? For a particular act to be considered rational, will it not be sufficient that it should be seen to be capable of maximizing an agent's long-range expectable utility, whether or not the expectable utility is from a dispassionate and informed point of view? Perhaps Richards may insist here that, if what a person desires is not what he also would desire from a dispassionate and informed point of view, his desire will not, as a matter of fact, maximize his long-range expectable utility.

This response does seem to provide a basis for distinguishing between real and apparent rational acts. A real rational act will be one that is desired under conditions of full information, and which is capable of maximizing an agent's "long-range" expectable utility. An apparent rational action, on the other hand, is one that does not in fact maximize the long-range expectable utility of the agent. Hence, to be truly or really rational, a person must not only seek to maximize his expectable utility. Rather, he must seek to maximize his long-range expectable utility. The idea of long-range thus becomes a very important constraint. Of any number of plans that can maximize a person's utility, that plan which is capable of maximizing his long-range expectable

utility is to be preferred. In a concrete example, we might say that if a person can both maximize his utility both by playing professional football and by practising medicine, if practising medicine maximizes his long-range expectable utility while playing football maximizes his short-term expectable utility, he will be irrational to sign up for a career of football. (Here we must assume that he has the required abilities to succeed in both careers.)

Now, there is nothing apparently odd with someone showing a preference for a career that would satisfy his long-range interests instead of a career that can only serve his short-term interests. The whole emphasis of secondary school counselling today is to help the student to articulate that plan of life that would best fulfill his long-range interests (life goals). However, is it so evident that such an appeal is value-free? The idea of long-range is obviously distinct from that of efficiency or effectiveness. Consider two alternative plans A and B. A is more efficient than B, but B promotes the agent's long-range utility, while A only promotes his short-term utility. According to the received view, B obviously is to be preferred. But is this not in effect saying that an agent's long-range utility is to be preferred to his short-term utility? And if this is so, are there no value presuppositions in such a claim?

The unique merit of Nielsen's principle lies in the fact that it articulates the normative (value) element implicit in our ordinary conception of rationality. It is this element that distinguishes the tenth principle from the fifth principle in his index. This element is well illustrated in his (25) commentary on Brandt's "Rationality, Egoism and Morality". (26) With Brandt's calculative model of rationality in mind, Nielsen develops a counter example as follows: we suppose someone arriving in a strange country for a day or so. On arrival, he notices that a movie he had so intensely desired to see is playing that night. As there is such a long queue of people ahead of him waiting to see the same movie, he is left with two options: cheat his way through the crowd or take his turn at the bottom of the queue with a risk of not seeing the movie. Should the stranger appeal solely to Brandt's calculative model, the most rational thing for him to do, it would seem is to try to cheat his way skillfully through the crowd. While such a course of action need not be seen as placing a great obex on his conscience, it is not plainly evident that he would be any less rational should he decide to take his proper place at the bottom of the line.

25 The fifth principle in Nielsen's index reads: "The most efficient and effective means are to be taken, ceteris paribus, to achieve one's ends". op. cit., p.57.

26 Kai Nielsen, "Ethical Egoism and Rational Action", Journal of Phil., Dec., 28, 1972, pp.698-700.

But from the point of view of Brandt's thesis, the latter choice can hardly be rational. Insofar as the stranger knows that his long-range self-interest would be best served by successfully beating the line, then the rational thing for him to do is to go ahead and beat it. Brandt cannot appeal to the constraint of fairness or morality since all that is required for his agent to be rational is the maximization of his long-range utility.⁽²⁷⁾

But is it not the case that, if a person is indeed rational, he would be concerned to see the members of his society committed to principled and fair behaviour? If a person is rational, he would see that his life as well as the life of everyone else in the society would be less pleasant if the condition of genuine fair behaviour does not exist in the society. This fact can adequately motivate his regard for fair behaviour even if he does not take it as mandatory on everyone including himself. (He might be a classist amoralist.)

27 Here we might also think of Richards' assertion that "acts may be right and moral, though irrational ... and rational, though immoral (e.g. white South Africans' creation of apartheid, which may advance their aims quite well, but which is none the less immoral). The concepts of rationality and morality simply have these different intensions and consequent different possible extensions, for the principles of rationality depend on the notion of best, or satisfactory, realizing the agent's aims, whatever they are, whereas the principles of morality put certain sorts of substantive constraint on the kinds of ends an agent may legitimately pursue". op. cit., pp.229-30. If the claims here are true, on what basis can someone still undertake a rational assessment of the

However, if the notion of "a rational man" only conveys (à la Brandt, Rawls and Richards), the idea of a prudent maximizer -- "a fully informed man" who knows his most valued desires and the best ways to realize them -- then it does not seem to follow that a rational man will necessarily be concerned to be fair. The most that can be said of him is that he will be a man of good morals. Of course, it cannot be taken simply for granted that a rational man will necessarily be a man of principled behaviour, even if he sees it as only having great instrumental value. But such a minimal guarantee is not even assured given the supposed morally neutral characterization of rationality. A rationality that is alleged to be value-free cannot lead a person to see the value of principled behaviour.

Thus, for an agent's desires and ends to be considered dispassionate and fully informed, the agent must be aware of what these desires and ends are; he must be aware of the possibility of attaining them and the most effective means of attaining them, and the moral implications of these desires and ends both from his standpoint and that of other members

desirability or non-desirability of an agent's ends? Are we then to take it that any end at all is desirable provided an agent has an effective means of attaining it? Also, if there is such a gulf between the notions of Morality and Rationality, how is the reconciliation of both concepts possible?

of his society. These are necessary criteria for applying the term "rational" to persons in the context of their ends and means as moral agents.

It is very important to note that the concept of rationality involves not only a choice of appropriate means to desired ends, but a choice of ends as well. When it is used to describe the choice of means then the question of coherence or consistence becomes very crucial. But when it is directed to the choice of ends, an appeal must be made to some value. And, as it applies to moral agents, rationality cannot be divorced from the end that the moral person seeks to attain. This end is the attainment of moral goodness for which a moral person necessarily acts when he rationally acts morally. If a moral agent is to make a rational choice therefore, the means as well as the end he seeks to realize must be consistent with the constraints of morality.

Thus, the richer conception of rationality that is being suggested here is one that includes some of the ideals that are commonly associated with the general requirements of the moral point of view. Ideals which we would ordinarily accept on reflection as constituting a sufficient basis for rational decisions. It is a conception of rationality that allows a person to act with a direct reason to promote the interests and well-being of the other person. This conception is more effective, and perhaps superior to that of Rawls'

because it is peculiar to moral theory and distinctive of morality. As noted earlier, the conception adopted by Rawls is not exactly value-free in that it requires an agent to show preference for his long-range interests as opposed to his short-term interests. Apart from the fact that this requirement does not fall strictly within the idea of "taking the most effective means toward a desired end", it is one that people might want to challenge. Even though it might be rational to promote our long-range interests most of the time, there are situations in which a person can act with a preference for his short-term interests and he cannot obviously be described as irrational. Consider the habit of smoking. Several medical researches have shown the hazards of smoking. In spite of this, it is hardly proper to describe everyone seen smoking a cigarette as being irrational.

Besides, the idea of "long-range" does not adequately discriminate between ends such that an agent can plainly see which of his ends are rational or morally acceptable. Consider a person who voluntarily desires the life plan of a slave. Whatever he does that would promote the long-range utility of a slave must be considered rational from the standpoint of the received view. But one can ask meaningfully and without contradiction whether a long-range desire or utility of a slave is, as a matter of fact, a rational thing. Perhaps, one might respond here that if the person had full information

concerning the notion of 'slave', he would not, as a matter of fact, desire the long-range utility of a slave. This might be so. It is doubtful, however, whether such full information can be free of value presuppositions.

It is rather difficult, if not impossible, for anyone to evolve a conception of rationality that is entirely free of value presuppositions. Hence, the real issue between the conception adopted by Rawls and the one that is being suggested here is not whether either of them is value free. The issue rather is: which value assumption is more widely acceptable? Rawls obviously would want to claim that the assumption made by his conception is more widely acceptable. Again, this might just be another claim that could be checked empirically. Assuming, however, that this is actually the case, the wide acceptance of this conception is not due to the fact that it is a more rational conception or that it is more distinctive of ideal human behaviour. It is more widely acceptable because of the set of values from which it emerges. Economic facts and values are generally less controversial than moral ideals. However, if Rawls is indeed interested in a conception of rationality that is peculiar to moral theory, he would need a conception that includes, at least, the values of morality and not just those of economics: a rationality that characterizes not just an economic agent, but a moral agent as well. Such a conception of rationality could also

be widely accepted provided it is seen from the standpoint of the moral person.

The moral sceptic might ask at this point why he ought to be "rational" in this richer sense if some of its features are not features of rationality as such. It is not very clear how a satisfactory answer could be given to such a question, if one is possible at all. The question seems quite intelligible and non-trivial if the "ought" is seen as a non-moral ought. As suggested earlier, the topic of rationality (its precise features) calls for a careful and indepth study. However, it might be useful to remind the sceptic of one important thing. That just as it seems impossible to show him a conclusive reason, why he ought to be rational in a moral sense, if he is not already committed to the moral point of view, it is equally impossible to provide a conclusive reason why he should be rational in an economic sense, if he does not want the ideals of economics. It is quite intelligible and meaningful for a person to ask why he ought to promote his long-range interests as opposed to the short-term ones. If he attaches more value to intensity, for instance, he could rationally opt to satisfy those short-term interests that yield intense enjoyment. Recall Aristotle's saying that a shorter span of a well-spent life is more desirable than a greater span of a less fulfilling life. To conclude our discussion, we shall briefly examine the second question raised earlier.

THE POSSIBILITY OF A CONTRACT

As stated earlier, there are at least two obvious reasons why the conception of rationality that was suggested in the preceding section cannot work within the present framework of Rawls' theory. First, the sort of difficulties surrounding this conception are some of the very things that Rawls' approach is designed to avoid. For instance, even though Rawls would not equate the problem of moral philosophy with that of ordinary rational choice, his attempt to convert the former into the latter is only a way to provide a less controversial explanation for the former. He can hardly make any significant headway if he were to assume a problematic conception of rationality in an attempt to provide a non-controversial account of moral choice.

Another important consideration is that the relevance of Rawls' original position would seem to raise serious question should he introduce any substantive moral concepts into it. In constructing the original position, Rawls wants to be able to derive moral principles from morally neutral rational principles. He can hardly attain this objective should he allow a conception of rationality with built-in moral ideals. In fact, if he were to proceed in this direction, it would appear that he could only determine what rational principles are partly in terms of antecedently understood moral conceptions. Even if such a move can be defended

against the charge of circularity, one might still ask why we even need the original position or the social contract at all. Besides, if Rawls hopes to persuade the amoralist or the moral sceptic to adopt the moral point of view, but if his conception of rationality already includes moral ideals, why should either of these persons take him serious? Why could he not consistently or rationally desire not to be rational from the point of view of morality? He might still be rational in an economic sense (he might still be able to maximize his long-range expectable utility) even if he is irrational through and through from the point of view of morality. While the enriched conception of rationality that was suggested above might enable us to identify what is rational from the moral point of view, we cannot proceed to apply it directly in every situation where the rationality of an agent is in question. We cannot regard any ordinary person as irrational just because his actions do not conform with this conception of rationality.

Thus there are at least two prima facie reasons why Rawls cannot react positively to the revision being advanced in this chapter. But does he really have much of a choice one way or the other? To be sure, if it is possible to provide a non-controversial and non-problematic account of moral choice, there is no reason why such an account should not be preferred. But has Rawls given us any reason to believe

that a non-controversial or non-problematic account of moral choice is possible? His account at least does not show this. If the objections which have been raised in this study are valid, it is doubtful whether the concepts employed by him are any less problematic than the ones being proposed here. Indeed, contrary to his claim and belief, the assumptions at the original position are by no means weak. If anything, they are just a set of different assumptions about human beings which must be argued for. For example, it is just as controversial to claim that human beings are primarily self-interested as it is to claim that they are altruistic. Thus, it might be useful and more beneficial to pay less attention to the controversial aspect of the conception of rationality being suggested, and concentrate rather on the substantive results it yields. It suffices to note that if two concepts are equally controversial, but one is able to explain a given phenomenon better than the other, then it is only reasonable to appeal to the one that explains better. Naturally Rawls would claim that his conception of rationality is not only less controversial, but that it explains better. However, if the arguments adduced in the preceding chapters are taken into account then Rawls would still need to do much more to establish his case. Again, the economic theory of rationality may indeed be less controversial than the alternative conception being suggested here. But the latter

cannot explain the moral phenomenon for the simple reason that a rational choice in terms of economics is essentially different from a rational choice in terms of the moral point of view.

Also, the attempt to characterize rationality in terms of moral conceptions may appear less appealing at first sight. But, here again, Rawls does not seem to have much of a choice. Rationality by its very nature cannot be characterized in isolation from a particular point of view. The very idea of a neutral conception of rationality is very contestable. Any conception of rationality necessarily presupposes some type of value. Now, because Rawls wants the features of his original position to be peculiar to moral theory, a rationality that is characterized with reference to an economic ideal cannot be appropriate. He needs a conception that takes moral ideals into account. This need not be considered as a weakness of any significance. Just as it would indeed be the case that an antecedent knowledge of moral ideals is required before the principles of such a rationality can be determined, it is also the case that some antecedent idea of economic ideal is required to be able to determine the principles of economic rationality. At any rate, this is an issue that requires further investigation.

In the light of these considerations therefore, it is important to ask whether a contract situation leading to

the derivation of the principles of justice does still exist if we are to implement the revision being proposed in this chapter. In other words, can the two principles of justice as fairness still emerge from an original position that includes a stronger or morally built-in conception of rationality?

This question can obviously be answered in the affirmative, although the relevance or significance of such a contract would still remain doubtful. In the first place, Rawls, in setting up the features of the original position, indicates that he wants its conditions to convey the idea of a well-ordered society: a society in which individuals see themselves as moral persons. One of the conditions of being considered a moral person, as he conceives it, is that a person should have a sense of justice that is normally effective. That is, a person must have the desire to act on the principles of justice, and these principles must determine his conduct most of the time. Now if an original position whose features suggest the idea of moral persons can generate principles of justice, and if a morally built-in conception of rationality is compatible with the notion of moral persons then it cannot impede the derivation of the principles of justice as such. Of course, there is still the question whether such an original position would lead to the kind of principles Rawls wants to derive. Again, one needs to be cautious here. If Rawls only wants to appeal to some

particular conception of a moral person in order to be able to derive a particular conception of justice, then he could easily be vulnerable to the charge of circularity and indeed a grave circularity at that. Thus, he cannot just exclude a particular assumption that is true of moral persons simply because it is incompatible with the particular notion of justice he wants to derive. Indeed, if there is an assumption that is true of moral persons as such, and such an assumption is incompatible with his conception of justice so much the worse for his conception.

Again, one might respond to this by drawing attention to the fact of the difference between the original position and the actual life conditions of a well-ordered society. Rawls has repeatedly drawn attention to this distinction in response to some of his critics. But how valid or adequate is such an appeal? It was argued earlier that ordinarily we might just consider the original position as a mere hypothetical device as Rawls himself designates it. In which case, we need not demand that the assumptions in it be correct or exactly descriptive of moral persons. But, as seen earlier, Rawls does not consider his original position as a mere hypothetical device in the sense that its assumptions are arbitrarily drawn. He assures us that his assumptions would be seen to be essentially correct when we must have carefully reflected on the nature of

morality and moral persons. The thin theory must be sound, he says. Also, Rawls insists that any choice of principles made in the light of these assumptions is bound to be a rational choice since the agent would always be satisfied, even when he has come to real life situation, that he made the best and most rational choice in the original position. He cannot regret his choice even when the veil of ignorance has been lifted. In other words, if he has to choose all over again, he would choose the same way as he had done in the original position. The original position is interpreted in (28) such a way "so that one can at any time adopt its perspective". If this is Rawls' attitude, then the substantive claims he makes at the original position must indeed be well-founded. Above all, his assumptions must meet the criteria of the notion of moral persons, and the presence of even one false premise in the original position would be sufficient to destroy its validity.

Thus, the principal issue here is whether the proposal put forward in this chapter is, as a matter of fact, compatible with the notion of moral persons. Is it indeed the case that in their choice of moral actions, moral persons, regardless of the particular conception of morality they happen to hold, would act in accordance with the principles of a rationality

that includes moral ideals? Again, this question can only be answered in the affirmative. It is doubtful whether Rawls himself can claim the contrary. All his vigorous attempt to establish the morality of principles in his account of the psychology of moral development aims precisely at this. When a person has reached the realm of the morality of principles, he sees the need to act for the sake of these principles. In other words, he would accept considerations of morality as a final reason for acting.⁽²⁹⁾

Therefore, a rationality that includes moral ideals cannot be considered incompatible with the original position to the extent that the original position is characterized in order to reflect the idea of persons with a sense of morality (sense of justice). But does a contract situation yet arise in such a situation, and what is the relevance of such an original position in the role of testing moral theories?

29 Again, we must keep in mind that Rawls' original position does not indicate how such a morality is possible. He has not been able to argue convincingly why mutually self-interested persons would become so motivated even when they become members of a well-ordered society. If they are only concerned to further their own individual interests in the original position, it would only be a matter of chance or luck should they prove otherwise in real society. Rawls' appeal to Piaget's Psychological reductionism does not work. Self-interested persons cannot even understand, let alone see, how a person can act for the sake of moral principles (that is, seeing the requirement of morality as final reasons for acting).

The first part of the question can also be answered in the affirmative. Contrary to Rawls' claim, a situation requiring a contract may yet arise even if the contractors in the original position are presumed to be rational in a moral sense and have a morally built-in conception of the good. A contract situation would still arise if one takes into account the Humean conditions of moderate scarcity. Even if the parties are presumed to be altruistic, this does not necessarily imply that they must have general affection for one another. Altruism, as characterized by Thomas Nagel is only an objective recognition of the reality of other persons who also have objective interests to be met. It is a recognition of the fact that an individual is just one person among others and that other persons have just as much right as himself.⁽³⁰⁾ Thus, even if Rawls' contractors are said to be altruists, there is still a need for them to come into agreement on a formula for sharing their common but limited resources fairly. An example might well illustrate this claim. Suppose there is a rectory which housed twenty Dominican priests, with one altar for saying mass. Even though these Dominican priests may be presumed to be men of genuine concern for principled behaviour, there is still a need for them to come to an agreement on how

³⁰ Thomas Nagel, The Possibility of Altruism, Clarendon Press, Oxford, (1970), p.3.

each one could gain access to the use of the altar for their private masses. The difference, of course, is that unlike Rawls' self-interested men, the agreement that these priests would arrive at what could properly be called a moral agreement. Thus, the condition of rational altruism does not destroy a contractual framework as such. If anything, it merely guarantees that the agents of the original position would be moral agents, capable of making moral judgments. And this happens to be what Rawls is looking for.

Now whether Rawls can still use the original position as a critical device or a device for testing the adequacy of moral theories is not so evident. If the original position already contains moral assumptions, such a move would have no significant merit even if it is possible. But this fact does not count against the inclusion of moral assumptions in the original position. Rather, it challenges the fundamental assumption that an original position that is morally neutral, or at least neutral to the extent that Rawls makes it, can provide an effective way for testing the adequacy of moral theories. Rawls has not successfully shown why such a device ought to be adequate. Thus, even if an introduction of moral assumptions would render the device ineffective, this would hardly be of any serious consequence since the device itself is doomed to fail even without the moral notions. The problem of moral choice is quite distinct from

the problem of rational prudential choice. To the extent that the original position is designed to cater for the latter, it cannot be successfully used for the latter.

This study will now be brought to a conclusion with a brief remark on the central lesson to be learnt from Rawls.

CONCLUSION: A LESSON FROM THE PLIGHT OF RAWLS' THEORY

The main features of Rawls' account have been kept constantly in sight throughout this study. While following his arguments wherever they led, summary statements showing the main discontents with these arguments have been provided at the conclusion of each chapter. It will therefore not be necessary to restate those objections here. Instead, these concluding remarks are intended to be a general commentary on Rawls' theory. Emphasis will be on the reason why Rawls' account is plagued with serious difficulties and the general lesson which contemporary writers in ethics might still learn from it.

The focus of the inquiry has been on the individual arguments which Rawls puts forward for the various aspects of his theory. An attempt has been made to point out the inherent difficulties in these arguments. But the matter can be pursued further. For instance, the question may be asked: why is Rawls' account so plagued with difficulties and could he avoid them? To answer a question such as this, one would have to examine the foundations on which Rawls' theory has been built. Thus in what follows, I shall attempt to specify some of the main reasons why Rawls' theory does not accomplish what it had promised.

One of the most important reasons for the apparent collapse of Rawls' theory is that he failed to recognize, or at

least to sufficiently take to heart and consider, that there is something peculiarly puzzling and problematic about the entire phenomenon of morals. This puzzle is not just about what is morally right or wrong (good or bad) but what it is for something to be morally right or wrong. Rawls' failure to recognize the importance of this question is rather obvious from what emerges as one of the foundational ideas of his theory. He wants, as he says, to convert the problem of moral choice to that of a theory of rational choice. The idea that such a conversion is possible suggests that the activity of ordinary rational (non-moral) choice and that of a moral decision are essentially the same, or that there is an intimate connection between the two. But is this so obvious the case? Can there not be quite ordinary rational people who have no moral views, or at least no moral commitments at all, and to whom morality means nothing? If there are such people, does this not indicate that the ground on which the consciousness of moral distinctions rests is quite different from that of ordinary non-moral choice?

The efficacy of Rawls' architectonic plan is not here in question. Moreover, quite a number of disciplines in the social and behavioural sciences would easily find certain elements in his theory agreeable. Nonetheless, the question still remains: does Rawls really have an adequate conception of what it is for something to be morally right or wrong? If he does, does his account reflect such an understanding?

Here we might think of a very important paradox, stressed by Aristotle in the earlier part of the Nicomachean Ethics. The paradox can be stated in the form of a question: how can a person do morally good acts if he is not already a morally good person? Whereas the performance of certain types of acts in accordance with some commonly accepted rules is sufficient to identify what type of art a person is engaged in, the same procedure does not follow so rigidly in the case of a moral act or a morally good (virtuous) person. The mere performance of certain types of acts does not make a person morally good or virtuous, even if such acts are known to be morally good. To be commended as a morally good or virtuous person, a person's action must, among other things, be motivated by morally good desires or reasons. A morally good person is one who desires to perform morally good acts for the sake of promoting moral goodness.

It is useful to think here too of an important criticism of Kant by Sidgwick. Put very simply the point is this: we need to distinguish clearly between questions of blame and praise where a person's motives are absolutely crucial and the distinction between a man of good morals and a morally good man is crucial. But a person may do what is in fact right (from some moral point of view) and do it from a morally despicable motive. For example someone keeps his promise hoping, by doing this, to curry favour with the person he keeps his promise to,

when it is clear that he would not do it but for the fact that he thinks by doing it he would succeed in getting in the favour of that person. What he did was right (it was the right thing to do and in that sense the person's act was morally good, that is, it would have been wrong for him to do otherwise) but he does not deserve praise for it. If indeed his motives were known he should be criticized. Similarly someone may do something which is wrong for morally praiseworthy motives. A gives B, who is wild with thirst, a drink of water thinking the water was clean when in fact it is not, and B dies from the drink. A did not do the right thing but he is blameless (if he had taken reasonable precautions) and his motives were good. As Sidgwick argues against Kant these are two different dimensions of moral assessment. We should make a distinction between right and wrong acts (morally good and bad acts) and good and bad motives.

At crucial points in his theory, Rawls misses the distinction (as Kantian as he is) between a man of good morals and a morally good man. His attempt to convert the problem of moral choice to that of non-moral rational choice cannot be a move in the right direction. A moral choice has more to it than mere internal coherence or consistency within a person's preferred life plan. Thus, even though Rawls has expounded a very sophisticated account of rational choice the entire project appears to be doomed to failure from the start. Assumptions based on

economic theory, game theory, psychology and the like are not the right tools for establishing the motivation to be morally good or just. These assumptions may have proved helpful in the sort of rigorous deductive formula that Rawls wants to generate, but the question is whether a normative theory of justice or morality can be susceptible to a neat and tidy deduction. If the entire field of morality is such a murky and problematic discipline, in what way is a rigorous deductive approach an asset? Might such an approach not even conceal or undermine some of the crucial features of morality? Again, Aristotle has rightly warned us not to expect that the principles of morality can be deductively established like the principles of logic. There is a distinction between what it is for something to be morally right or wrong, and for something to be in accordance with the rules of valid inference. A principle of action may be reasonable and acceptable from the moral point of view, even if it is unjustifiable from the standpoint of economic theory or otherwise.

Rawls' theory of the good does make a strong initial appeal. His account seems to recognize an important fact: that even though the term "good" is used of many different sorts of things, of things indeed in different categories, nonetheless, the term is not that mysterious or ambiguous. There is still some degree of consistency in the various senses or contexts in which it is used. For instance, although the good making

properties of a military general are quite different from those of a strawberry, there is still something common and consistent in the two expressions: "a good general" and "a good strawberry". The good making properties in both expressions have a basis in fact. They are the qualities that it is rational to desire from objects of their kind.

Rawls' account also captures a very important logical feature of the term "good". In many of its occurrences, the term "good" functions as an attributive and not as a predicative adjective. For example the adjective "green" is predicative. Consider a statement: "this is a green snake." This statement can be further analyzed as "this is a snake and it is green." And since a snake is a reptile, we can further conclude that "this is a green reptile." But the term "good" does not function in such a predicative sense. We cannot derive the statement "this is a good man" from the statement "this is a good footballer." The adjective "good" is usually so to say, 'glued' to the substantive it qualifies. It is in view of this logical behaviour that it can be called an attributive adjective. Rawls' account articulates this characteristic of good. In his account the "good" in "good X" is attributive because "good X" means something like "better than most of average Xs". It is an X that has the properties (to a higher degree than the average or standard X) which it is rational to want in an X, given what X's are used for, or expected to do, and the like.

However, Rawls' account begins to run into serious difficulties the moment he moves from the goodness of roles and artifacts to the goodness of persons and moral standards. What Rawls does in the latter case is to develop his thin theory into the full theory of the good via the original position. This consists in applying the thin theory together with the two principles emerging from the original position to the two basic cases of a good person and a good society. Hence, as he says, a good person is one who has to a higher degree than the average person the broadly based properties that it is rational for persons to want in one another in a well-ordered society. These properties include a sense of justice. To be considered as good persons therefore, members of a well-ordered society must affirm their sense of justice and desire other members of their society to do likewise.

One important consideration in favour of this definition is that it effectively undermines, or so it seems, the question whether or not there is a role or purpose for which man is created. By introducing the notion of well-ordered society Rawls is able to assert the idea of a shared common rational end. Members of a well-ordered society have a shared common end of maintaining their just institutions. To realize this end, they must have, and desire that others should have, an effective sense of justice.

All this is a castle built on sand, since Rawls does not have an adequate theory of how an effective sense of justice arises in the first instance. The motivational assumptions together with the model of rationality in his original position cannot yield an effective sense of justice. Indeed the members of Rawls' well-ordered society may not be any more than rational prudent maximizers of a Hobbesian society. They might be men of good morals but not morally good persons.

As regards the challenge of moral scepticism, it does not appear that Rawls' account has been able to undermine this challenge in any important way. His non-moral conception of rationality cannot sufficiently induce an amoralist to endorse the moral point of view. His moral epistemology does not show how we can attain the type of objectivity the ethical subjectivist is arguing for. He has not succeeded in showing that there are moral judgments whose truth is publicly determinable by any rational observer who is acquainted with the facts (moral judgments that are valid for all rational persons). It is doubtful of course, whether such an objective moral theory can be constructed. This is an issue that requires careful investigation. Also the notion of rationality as it applies in the context of morality needs further study. And, until such investigations have been conducted and some positive conclusions reached, any claim of objectivity in ethics must be pushed with caution.

BIBLIOGRAPHY

- Ake, Christopher, "Justice as Equality," Philosophy and Public Affairs, 5(Fall, 1975), 69-89.
- Albert, Ethel M., Demise, T. C., and Peterfreund, S. P. (eds.), Great Traditions in Ethics, 3rd ed. (New York: D. Van Nostrand Co., 1975), Chap. 19, 416-37.
- Alexander, S. S., "Social Evaluation Through Notional Choice," The Quarterly Journal of Economics, 88(1974), 597-624.
- Altham, J. E. J., "Rawls' Difference Principle," Philosophy, 48(1973), 75-8.
- Andelson, Robert V., "Vive La Difference? Rawls' Difference Principle" and the Fatal Premise Upon Which It Rests," The Personalist, 56(1975), 207-13.
- Arrow, Kenneth J., Social Choice and Individual Values, second edition (New York: John Wiley and Sons, 1963).
- _____, "The Utilitarian Approach to the Concept of Equality in Public Expenditures," The Quarterly Journal of Economics, 85(1971), 409-15. Also in Economic Justice, ed. by E. S. Phelps (Baltimore: Penguin Books, 1973), 439-46.
- _____, "Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice," The Journal of Philosophy, 70(May 10, 1973), 245-63.
- Atkinson, A. B., "How Progressive Should Income-Tax Be?" in Essays on Modern Economics, ed. by M. Parkin (New York: Longmans, Green and Co., Inc.). Also in Economic Justice, ed. by E. S. Phelps (Baltimore: Penguin Books, 1973), 386-404.
- Baier, Kurt, "Moral Development," The Monist, 58(1974), 601-615.
- _____; "Individual Moral Development and Social Moral Advance," a symposium with Richard G. Henson and Lawrence Kohlberg, The Journal of Philosophy, 70 (October 25, 1973), 646-8.

- Barber, Benjamin R., "Justifying Justice: Problems of Psychology, Measurement, and Politics in Rawls," The American Political Science Review, 69(1975), 663-74. Also in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 292-318.
- Barry, Brian, "On Social Justice," The Oxford Review (Trinity Term, 1967), 33-43. Also in Justice and Equality, ed. by H. Bedau (Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1971), 103-15.
- _____, "John Rawls and the Priority of Liberty," Philosophy and Public Affairs, 2(Spring, 1973), 274-90.
- _____, The Liberal Theory of Justice, (London: Oxford University Press, 1973).
- _____, "Liberalism and Want-Satisfaction: A Critique of John Rawls," Political Theory, 1(1973), 134-53.
- Bates, Stanley, "The Motivation to be Just," Ethics, 85 (1974), 1-17.
- Beardsley, Monroe, "Equality and Obedience to Law," in Law and Philosophy, ed. by Sidney Hook (New York: New York University Press, 1964), 35-42.
- Becker, Edward F., "Justice, Utility, and Interpersonal Comparisons," Theory and Decision, 6(November, 1975), 471-84.
- Bedau, Hugo A., "Review of Brian Barry's The Liberal Theory of Justice," The Philosophical Review, 84(1975), 598-603.
- Beitz, Charles R., "Justice and International Relations," Philosophy and Public Affairs, 4(Summer, 1975), 360-89.
- Bentley, D. J., "John Rawls: A Theory of Justice," The University of Pennsylvania Law Review, 121(1972-73) 1070-78.
- Bierman, A. K., "Chessing Around," Philosophical Studies, 23(1972), 141-2.
- Bloom, Allan, "Justice: John Rawls vs. The Tradition of Political Philosophy," The American Political Science Review, 69(1975), 648-62.

Bowie, Norman, Towards a New Theory of Distributive Justice (Amherst, Mass.: The University of Massachusetts Press 1971).

_____, "Some Comments on Rawls' Theory of Justice," Social Theory and Practice, 3(1974), 65-74.

Brandt, Richard C., "Utility and the Obligation to Obey the Law," in Law and Philosophy, ed. by Sidney Hook (New York: New York University Press, 1964), 43-55.

Braybrooke, David, "Utilitarianism with a Difference: Rawls's Position in Ethics," The Canadian Journal of Philosophy, 3(1973), 303-31.

Brock, Dan W., "Contractualism, Utilitarianism and Social Inequalities," Social Theory and Practice, 1(1971), 33-44.

_____, "Recent Work in Utilitarianism," The American Philosophical Quarterly, 10(1973), 241-76.

_____, "John Rawls's Theory of Justice," a symposium with Ronald Dworkin and H. L. A. Hart, The University of Chicago Law Review, 40(1973), 486-99.

Brown, D. G., "John Rawls: John Mill," Dialogue (Canada), 12(1973), 477-9.

Buchanan, Allen, "Distributive Justice and Legitimate Expectations," Philosophical Studies, 28(December, 1975), 419-25.

Buchanan, James M., "Boundaries on Social Contract," Reason Papers, 2(Fall, 1975), 15-28.

_____ and Tullock, G., The Calculus of Consent: Logical Foundations of Constitutional Democracy (Ann Arbor, Michigan: The University of Michigan Press, 1962).

_____ and Bush, W. C., "Political Constraints on Contractual Redistribution," a symposium on Concepts of Distributional Equity in The American Economic Review, 64(May, 1974), 153-7.

Burkholder, L., "Rule-Utilitarianism and 'Two Concepts of Rules'," The Personalist, 56(1975), 195-98.

- Campbell, Richmond, "Review of Rawls's A Theory of Justice," Dalhousie Law Journal, 1(1973), 210-23.
- Campbell, T. D., "Humanity Before Justice," The British Journal of Political Science, 4(1974), 1-16.
- Card, Claudia, "Retributive Penal Liability," in Studies in Ethics: The American Philosophical Quarterly Monograph No. 7 (1973), 17-35.
- Care, Norman S., "Runciman on Social Inequality," The Philosophical Quarterly, 18(1968), 151-4.
- _____, "Contractualism and Moral Criticism," The Review of Metaphysics, 23(1969), 85-101.
- _____, "Review of Brian Barry's The Liberal Theory of Justice," Mind, 85(1976), 126-8.
- Carr, Spencer, "Rawls, Contractarianism, and our Moral Intuitions," The Personalist, 56(1975), 83-95.
- Chapman, John W., "Justice and Fairness," in Nomos VI: Justice, ed. by C. J. Friedrich and John W. Chapman (New York: The Atherton Press, 1963), 147-69.
- _____, "Natural Rights and Justice in Liberalism," in Political Theory and the Rights of Man, ed. by D. D. Raphael (Bloomington, Indiana: Indiana University Press, 1967), 27-42.
- _____, "Rawls's Theory of Justice," The American Political Science Review, 69(1975), 588-93.
- Charvet, John, "The Idea of Equality as a Substantive Principle of Society," Political Studies, 17(1969), 1-13.
- Chevigny, P. G., "A Theory of Justice," Civil Liberties, 298(September, 1973), 4-5.
- Choptiany, Leonard, "A Critique of John Rawls's Principles of Justice," Ethics, 83(1973), 146-50.
- Cohen, Marshall, "Review of A Theory of Justice," New York Times Book Review (July 16, 1972).

- Coleman, James S., "Beyond Pareto Optimality," in Philosophy, Science, and Method: Essays in Honor of Ernest Nagel, ed. by Morgenbesser, Sidney, Suppes, Patrick, and White, Morton (New York: St. Martin's Press, 1969), 415-39.
- Copp, David, "Justice and the Difference Principle," The Canadian Journal of Philosophy, 4(1974), 229-40.
- Craig, Leon H., "Contra Contract: A Brief Against John Rawls' Theory of Justice," The Canadian Journal of Political Science, 8(1975), 63-81.
- Cunningham, R. L., "Justice: Efficiency or Fairness?" The Personalist, 52(1971), 253-81.
- Dalgarno, M. T., "Review of Rawls' A Theory of Justice," Philosophical Books, 14(1973), 26-8.
- Daniels, Norman, "On Liberty and Inequality in Rawls," Social Theory and Practice, 3(1974), 149-59.
- _____ (ed.), Reading Rawls: Critical Studies of a Theory of Justice (New York: Basic Books Inc., 1975).
- _____, "Equal Liberty and Unequal Worth of Liberty," in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 253-81.
- Danielson, Peter, "Theories, Intuitions and the Problem of World-Wide Distributive Justice," a symposium with Nowell-Smith and C. B. Macpherson, Philosophy of the Social Sciences, 3(1973), 331-8.
- Darwall, Stephen, "A Defense of the Kantian Interpretation," Ethics, 86(No. 2, 1976), 164-70.
- Dasgupta, Partha, "On Some Problems Arising from Professor John Rawls' Conception of Distributive Justice," Theory and Decision, 4(1974), 325-44.
- DeMarco, Joseph P., "Some Problems in Rawls' theory of Justice," Philosophy in Context, 2(1973), 41-8.
- Dias, R. W. M., "Review of Rawls' A Theory of Justice," Cambridge Law Review, 30(1972), 353-5.

Dryer, D. P., "Review of Rawls' A Theory of Justice,"
The University of Toronto Law Journal, 5(1972), 249-51.

Dworkin, Gerald, "Non-Neutral Principles," in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 124-40. Also in The Journal of Philosophy, 71 (August 15, 1974), 491-506.

Dworkin, Ronald, "The Original Position," a symposium with Dan W. Brock and H. L. A. Hart, The University of Chicago Law Review, 40(1973), 500-533. Also in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 16-52.

Emmet, Dorothy and Morrison, Alasdair, "Justice," The Aristotelian Society Supplementary Volume, 40(1969), 109-40.

Eshete, Andreas, "Contractarianism and the Scope of Justice," Ethics, 85(1974), 38-49.

Feinberg, Joel, "Justice, Fairness and Rationality,"
Yale Law Journal, 81(1972), 1004-31.

_____, "Duty and Obligation in the Non-Ideal World," The Journal of Philosophy, 70(May 10, 1973), 263-75.

_____, Social Philosophy (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973).

_____, "Rawls and Intuitionism," in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 108-23.

Fishkin, James, "Justice and Rationality: Some Objections to the Central Argument in Rawls's Theory," The American Political Science Review, 69(1975), 615-29.

Fisk, Milton, "History and Reason in Rawls' Moral Theory," in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 53-80.

Flathman, Richard E., Political Obligation (London: Croom Helm Ltd., 1973).

- Flew, Antony, "Three Questions About Justice in Hume's Treatise," The Philosophical Quarterly, 26(January, 1976), 1-13. See especially Part III.
- Frankel, Charles, "Justice and Rationality," in Philosophy, Science, and Method: Essays in Honor of Ernest Nagel, ed. by Morgenbesser, Sidney, Suppes, Patrick, and White, Morton (New York: St. Martin's Press, 1969), 400-14.
- _____, "Review of A Theory of Justice," Columbia Human Rights Law Review, 5(1973), 547-64.
- _____, "The New Egalitarianism and the Old," Commentary, (September, 1973), 54-61.
- _____, "Justice, Utilitarianism, and Rights," Social Theory and Practice, 3(1974), 27-46.
- Fried, Charles, "Natural Law and the Concept of Justice," Ethics, 74(1964), 237-54.
- _____, "Review of Rawls' A Theory of Justice," Harvard Law Review, 85(1971-72), 1691-7.
- Friedmann, Wolfgang, "A Theory of Justice: A Lawyer's Critique," The Columbia Journal of Transnational Law, 11(1972), 369-79.
- Fuchs, Alan E., "The Concept of Morality in Rawls's Theory," an abstract, The Journal of Philosophy, 72(October 23, 1975), 6289.
- Gardner, Michael, "Rawls on the Maximin Rule and Distributive Justice," Philosophical Studies, 27(1975), 255-70.
- Gastil, Raymond D., "Beyond a Theory of Justice," Ethics, 85(1975), 183-94.
- Gauthier, David, "Justice and Natural Endowment: Toward a Critique of Rawls' Ideological Framework," Social Theory and Practice, 3(1974), 3-26.
- George, Lawrence C., "Review of A Theory of Justice," The University of California Los Angeles Law Review, 20(1972-73), 856-63.

- Gilbert, Joseph, "Interests, Role Reversal, Universalizability and the Principle of Mutual Acknowledgement," in Ethics and Social Justice, ed. by Kiefer, Howard E. and Munitz, Milton K. (Albany, New York: State University of New York Press, 1968), 164-70. A reply to Paul W. Taylor's article listed below.
- Gordon, Scott, "John Rawls's Difference Principle, Utilitarianism and the Optimum Degree of Inequality," The Journal of Philosophy, 70(May 10, 1973), 275-80.
- Gourevitch, Victor, "Rawls on Justice," The Review of Metaphysics, 28(1975), 485-519.
- Gow, H. B. "Defining the Just Society," The University Bookman, (Winter, 1973), 41-4.
- Grey, Thomas C., "The First Virtue," Standford Law Review, 25(1973), 286-327.
- Grice, Russell, The Grounds of Moral Judgment (London: Cambridge University Press, 1967).
- Grover, R. A., "The Ranking Assumption," Theory and Decision, 4(1974), 277-99.
- Haksar, Vitit, "Rawls' Theory of Justice," Analysis, 32.5(1972), 149-53.
- _____, "Autonomy, Justice and Contractarianism," The British Journal of Political Science, 3(1973), 487-509.
- Hall, Everett, "Justice as Fairness: A Modernized Version of the Social Contract," The Journal of Philosophy, 54(October 24, 1957), 662-70.
- Hampshire, Stuart, "A New Philosophy of the Just Society," The New York Review of Books, 3:3(February 24, 1972), 34-9.
- Hare, R. M., "Rawls' Theory of Justice, Part I, II," The Philosophical Quarterly, 23(1973), 144-55; 241-52. Also slightly revised in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 81-107.

- Harris, C. E., "Rawls on Justification in Ethics," The Southwestern Journal of Philosophy, 5(1974), 135-43.
- Hart, H. L. A., "Rawls on Liberty and Its Priority," a symposium with Dan W. Brock and Ronald Dworkin, The University of Chicago Law Review, 40(1973), 534-55. Also in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 230-52.
- Harsanyi, John C., "Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory," The American Political Science Review, 69(1975), 594-606.
- Henson, Richard G., "Correlativity and Reversibility," a symposium with Kurt Baier and Lawrence Kohlberg, The Journal of Philosophy, 70(October 25, 1973), 648-9.
- Hicks, Joe H., "Philosophers' Contracts and the Law," Ethics, 85(1974), 18-37.
- Holborow, Les, "Desert, Equality and Injustice," Philosophy, 50(1975), 157-68.
- Honderich, Ted, "The Use of the Basic Proposition of a Theory of Justice," Mind, 84(1975), 63-78.
- Hook, Sidney (ed.), Law and Philosophy (New York: New York University Press, 1964).
- _____, "Law, Justice, and Obedience," in Law and Philosophy, ed. by Sidney Hook (New York: New York University Press, 1964), 56-60.
- Hospers, John, "A Review of Rawls's A Theory of Justice," The Personalist, 55(1974), 71-7.
- Johnson, Conrad D., "Actual -V (Rawlsian) Hypothetical-Consent," Philosophical Studies, 28(1975), 41-8.
- Johnson, Oliver A., "The Kantian Interpretation," Ethics, 85(1974), 58-66.
- Johnson, Ronald, "Review of A Theory of Justice," Philosophy and Rhetoric, 8(1975), 133-5.
- Kava, Gregory, "Rawls on Average and Total Utility," Philosophical Studies, 27(1975), 237-53.

- Keat, Russell and Miller, David, "Understanding Justice," Political Theory, 2(1974), 3-31.
- Keyt, David, "The Social Contract as an Analytic, Justificatory and Polemic Device," The Canadian Journal of Philosophy, 4(1974), 241-52.
- Kiefer, Howard E. and Munitz, Milton K., Ethics and Social Justice (Albany, New York: State University of New York Press, 1968).
- Kleinberger, Aharon, F., "The Social-Contract strategy for the Justification of Moral Principles," Journal of Moral Education, Vol. 5, No. 2, (1976), 107-126.
- Klevatorick, Alvin K., "Discussion [of Rawls's Maximin Criterion]," The American Economic Review, 64(May, 1974), 158-61.
- Kohlberg, Lawrence, "The Claim to Moral Adequacy of the Highest Stage of Moral Judgement," a symposium with Kurt Baier and Richard G. Henson, The Journal of Philosophy, 70(October 25, 1973), 630-46..
- Konvitz, Milton, "Civil Disobedience and the Duty of Fair Play," in Law and Philosophy, ed. by Sidney Hook (New York: New York University Press, 1964), 19-28.
- Krislov, Samuel, "Justice as Formal Theory," Polity, 6(1974), 574-8.
- Ladd, John, "Law and Morality: Internalism vs. Externalism," in Law and Philosophy, ed. by Sidney Hook (New York: New York University Press, 1964), 61-71.
- Ladenson, Robert F., "Rawls' Principle of Equal Liberty," Philosophical Studies, 28(1975), 49-54.
- Lessnoff, Michael, "John Rawls' Theory of Justice," Political Studies, 19(1971), 63-80.
- _____, "Barry on Rawls' Priority of Liberty," Political Studies, 4(1974), 100-14.
- Levin, J., "Review of Rawls's A Theory of Justice," Modern Law Review, 36(1973), 667-70.
- Levine, Andrew, "Rawls' Kantianism," Social Theory and Practice, 3(1974), 47-63.

Lukes, Steven, Individualism (London: Basil Blackwell, Ltd., 1973).

_____, "Relativism: Cognitive and Moral," The Aristotelian Society, 48(1974), 165-89.

Lyons, David, Forms and Limits of Utilitarianism (London: Oxford University Press, 1965).

_____, "Rawls versus Utilitarianism," in "Symposium: A Theory of Justice," The Journal of Philosophy, 69(October 5, 1972), 535-45.

_____, "On Formal Justice," Cornell Law Review, 58(1973), 833-61.

_____, "The Nature of the Contract Argument," The Cornell Law Review, 59(1973-74), 1064-76.

_____, "Nature and Soundness of the Contract and Coherence Arguments," in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 141-67.

Mabe, Alan R., "Review of A Theory of Justice," The Journal of Politics, 35(1973), 1018-20.

MacCormick, Neil, "Justice According to Rawls," The Law Quarterly Review, 89(1973), 393-417.

MacIntyre, Alasdair, "Justice: A New Theory and Some Old Questions," Boston University Law Review, 52(1972), 330-4.

Mack, Eric, "Distributionism versus Justice," Ethics, 86(No. 2, 1976), 145-53.

Mackenzie, Nollaig, "An Alternative Definition of the Difference Principle," Dialogue: Canadian Philosophical Review, 13(1974), 787-93.

Macleod, A. M., "Critical Notice: Rawls' Theory of Justice," Dialogue: Canadian Philosophical Review, 13(1974), 139-59.

Macpherson, C. B., "Rawls' Models of Man and Society," a symposium with Nowell-Smith and Danielson, Philosophy of the Social Sciences, 3(1973), 341-7.

- Mandelbaum, M., "Review of A Theory of Justice," History and Theory, 12(1973), 240-50.
- Margolis, Joseph, "Justice as Fairness," The Humanist, 33(1973), 36-7.
- Masterson, M. P., "On Being Unfair to Rawls, Rousseau and Williams or John Charvet and the Incoherence of Inequality," The British Journal of Political Science, 1(1971), 209-222.
- Merritt, Gilbert, "Justice as Fairness: A Commentary on Rawls's New Theory of Justice," Vanderbilt Law Review, 26(1973), 665-86.
- Meyer, Michel, "The Perelman-Rawls Debate on Justice," Revue Internationale de Philosophie, 29(No. 113, 1975), 316-31.
- Michelman, Frank I., "In Pursuit of Constitutional Welfare Rights: One View of Rawls' Theory of Justice," The University of Pennsylvania Law Review, 121(1972-73), 962-1019. Also in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 319-47.
- Miller, Richard, "Rawls and Marxism," Philosophy and Public Affairs, 3(1974), 167-91. A revised version is in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 206-29.
- Miller, Richard W., "Rawls, Risk, and Utilitarianism," Philosophical Studies, 28(1975), 55-61.
- Morgenbesser, Sidney, Suppes, Patrick, and White, Morton (eds.), Philosophy, Science, and Method: Essays in Honor of Ernest Nagel (New York: St. Martin's Press, 1969).
- Morris, Bertram, "Rawls' Egalitarianism," Philosophical Research and Analysis, 5(1975), 2-3.
- Mortimore, G. W., "An Ideal of Equality," Mind, 77 (1968), 222-42.
- Mow, Joseph, "Rawls on Mercy: Pardon and Amnesty," The Journal of West Virginia Philosophical Society, (Spring, 1975), 2-5.

Mueller, Dennis C., "Achieving the Just Polity," The American Economic Review, 64(May, 1974), 147-52.

_____, "Intergenerational Justice and the Social Discount Rate," Theory and Decision, 5(1974), 263-73.

_____, Tollison, D. and Willit, T. D., "The Utilitarian Contract: A Generalization of Rawls' Theory of Justice," Theory and Decision, 4(1974), 345-67.

Murphy, Cornelius, "Distributive Justice, Its Modern Significance," The American Journal of Jurisprudence, 17(1972), 153-65.

Murray, John Courtney, Jr., "The Problem of Mr. Rawls's Problem," in Law and Philosophy, ed. by Sidney Hook (New York: New York University Press, 1964), 29-34.

Musgrave, R. A., "Maximin, Uncertainty, and the Leisure Trade-off," The Quarterly Journal of Economics, 88 (1974), 625-32.

McBride, William L., "Social Theory Sub Specie Aeternitatis: A New Perspective," Yale Law Journal, 81(1972), 980-1003.

_____, "The Concept of Justice in Marx, Engels, and Others," Ethics, 85(1975), 204-18.

McClennon, Edward, "Review of Brian Barry's The Liberal Theory of Justice," Social Theory and Practice, 3 (1974), 117-22.

Nagel, Ernest, "Fair Play and Civil Disobedience," in Law and Philosophy, ed. by Sidney Hook (New York: New York University Press, 1964), 72-6.

Nagel, Thomas, "Rawls on Justice," The Philosophical Review, 82(1973), 220-34. Also in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 1-15.

Narveson, Jan, Morality and Utility (Baltimore: Johns Hopkins Press, 1967), esp. 219-24.

Nathan, N. M. L., The Concept of Justice (London: The Macmillan Co., 1971), see the Appendix, "Rawls's Theory of Justice," 70-7.

Nelson, William, "Special Rights, General Rights, and Social Justice," Philosophy and Public Affairs, 3(1974), 411-30.

- Nielsen, Kai, "A Note on Rationality," The Journal of Critical Analysis, 4(1972), 16-9.
- _____, "A Critical Notice of John Rawls' A Theory of Justice," Second Order: An African Journal of Philosophy, Vol. 11, No. 2, (July 1973), 101-121.
- _____, "Does Ethical Subjectivism Have A Coherent Form," Philosophy and Phenomenological Research, Vol. 35, No. 1, (1974), 93-99.
- _____, "On Ascertaining What Is Intrinsically Good," The Journal of Value Inquiry, Vol. X, No. 2, (Summer 1976), 137-9.
- _____, "On Philosophic Method," International Philosophical Quarterly, Vol. XVI, No. 3, (September, 1976), 349-68.
- _____, "The Priority of Liberty Examined," The Indian Political Science Review, Vol. XI, No. 1, (January, 1977), 48-59.
- _____, "Rawls and Classist Amoralism," Mind, Vol. LXXXVI, No. 341, (January, 1977), 19-30.
- _____, "The Choice Between Perfectionism And Rawlsian Contractarianism," Interpretation: A Journal of Political Philosophy, Vol. 6, No. 2, (May, 1977), 132-9.
- _____, New Essays on Contract Theory, editors Kai Nielsen and Roger Shinner, (Guelph, Ontario: Canadian Ass. For Pub. in Philosophy, 1977).
- _____, "Class and Justice," Forthcoming in Justice And Economic Distribution, ed. John Arthur, (Englewoods Cliffs, N.J., Prentice-Hall), 1978.
- Nisbet, Robert, "The Pursuit of Equality," The Public Interest, (Fall, 1973), 103-20.
- Norton, David, "Rawls's Theory of Justice: A 'Perfectionist' Rejoinder," Ethics, 85(1974), 50-7.
- Nowell-Smith, P. H., "A Theory of Justice?" a symposium with Danielson and C. B. Macpherson, Philosophy of the Social Sciences, 3(1973), 315-39.

Nozick, Robert, "Moral Complications and Moral Structures," Natural Law Forum, 13(1969), 1-50.

_____, "Distributive Justice," Philosophy and Public Affairs, 3(Fall, 1973), 45-126.

_____, Anarchy, State, and Utopia (New York: Basic Books, Inc., 1974).

Oberdieck, Hans, "A Theory of Justice," New York University Law Review, 47(1972), 1012-28.

O'Connor, J., "Wolff, Rawls, and the Principle of Justice," Philosophical Studies, 19(1968), 93-5.

Parekh, B., "Reflections on Rawls' Theory of Justice," Political Studies, 20(1972), 479-83.

Pattanaik, P. K., "Risk, Impersonality, and the Social Welfare Function," The Journal of Political Economy, 76 (December, 1968), 1152-69. Also in Economic Justice, ed. by E. S. Phelps (Baltimore: Penguin Books Inc., 1973), 298-318.

Pennock, J. Roland, "The Obligation to Obey the Law and the Ends of the State," in Law and Philosophy, ed. by Sidney Hook (New York: New York University Press, 1964), 77-85.

Perelman, Chaim, Justice (New York: Random House Inc., 1967), esp. 39-52.

Pettit, Philip, "A Theory of Justice?" Theory and Decision, 4(1974), 311-24.

Phelps, E. S., (ed.), Economic Justice (Baltimore: Penguin Books Inc., 1973).

_____, "Taxation of Wage Income for Economic Justice," The Quarterly Journal of Economics, 87(1973), 331-54. Also in Economic Justice, ed. by E. S. Phelps (Baltimore: Penguin Books Inc., 1973), 417-38.

Pollock, Lansing, "A Dilemma for Rawls?" Philosophical Studies, 22(1971), 37-43.

Pritchard, M. S., "Human Dignity and Justice," Ethics, 82(1972), 299-313.

Proudfoot, Wayne, "Rawls on the Individual and the Social," The Journal of Religious Ethics, 2:2(1974), 107-28.

Rae, Douglas, "Maximin Justice and an Alternative Principle of General Advantage," The American Political Science Review, 69(1975), 630-47.

_____, "The Limits of Consensual Decision," The American Political Science Review, 69(December, 1975), 1270-94. See also Rae's rejoinder to Tullock's article below, Ibid., 1298.

Rand, Ayn, "An Untitled Letter," The Ayn Rand Newsletter, Vol. II, No. 10, 11.

Raphael, D. D., "Conservative and Prosthetic Justice," Political Studies, 12(1964), 149-62.

_____, "Critical Notice: Rawls' Theory of Justice," Mind, 83(1974), 118-27.

_____, "The Standard of Morals," Proceedings of the Aristotelian Society, 75(1974-75), 1-12; 12a-e.

Rasmussen, Douglas B., "A Critique of Rawls' Theory of Justice," The Personalist, 55(1974), 303-18.

Rawls, John, "Outline of a Decision Procedure for Ethics," The Philosophical Review, 60(1951), 177-97.

_____, "Two Concepts of Rules," The Philosophical Review, 64(1955), 3-32.

_____, "Justice as Fairness," The Journal of Philosophy, 54(October 24, 1957), 653-62.

_____, "Justice as Fairness," The Philosophical Review, 67(1958), 164-94.

_____, "Constitutional Liberty and the Concept of Justice," in Nomos VI: Justice, ed. by C. J. Friedrich and John W. Chapman (New York: The Atherton Press, 1963), 98-125.

_____, "The Sense of Justice," The Philosophical Review, 72(1963), 281-305.

_____, "Legal Obligation and the Duty of Fair Play," in Law and Philosophy, ed. by Sidney Hook (New York: New York University Press, 1964), 1-18.

_____, "Distributive Justice," in Philosophy, Politics and Society, Third Series, ed. by Peter Laslett and W. G. Runciman (London: Basil Blackwell Ltd., 1967), 58-82.

_____, "Distributive Justice: Some Addenda," Natural Law Forum, 13(1968), 51-71.

_____, "Distributive Justice," in Economic Justice, ed. by E. S. Phelps (Baltimore: Penguin Books Inc., 1973), 319-62. On p.319 Rawls says this is a restatement rather than a revision of his two previous articles on distributive justice.

_____, "The Justification of Civil Disobedience," in Revolution and the Rule of Law, ed. by Edward Kent (Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1971), 30-45.

_____, "Justice as Reciprocity," in Utilitarianism: John Stuart Mill, with Critical Essays, ed. by Samuel Gorovitz (New York: Bobbs-Merrill Co. Inc., 1971), 242-68.

_____, A Theory of Justice (Cambridge, Massachusetts: Harvard University Press, 1971).

_____, "Reply to Lyons and Teitelman," in "Symposium: A Theory of Justice," The Journal of Philosophy, 69(October 5, 1972), 556-7.

_____, "Some Reasons for the Maximin Criterion," a symposium on Concepts of Distributional Equity, The American Economic Review, 64(1974), 141-6.

_____, "Reply to Alexander and Musgrave," The Quarterly Journal of Economics, 88(1974), 633-55.

_____, "The Independence of Moral Theory," Presidential Address to the American Philosophical Association, Eastern Division, 1974, in Proceedings and Adresses of the American Philosophical Association, 48 (1975), 5-22.

_____, "Fairness to Goodness," The Philosophical Review, 84(1975), 536-54.

- Reiman, Jeffrey, "A Reply to Choptiany on Rawls on Justice," Ethics, 84(1974), 262-5.
- Rempel, H. D., "Justice as Efficiency," Ethics, 79(1969), 150-5.
- Rescher, Nicholas, Distributive Justice (New York: Bobbs-Merrill Co. Inc., 1966).
- Richards, D. A. J., A Theory of Reasons for Action (London: Oxford University Press, 1971).
- Ross, Geoffrey, "Utilities for Distributive Justice," Theory and Decision, 4(1974), 239-58.
- Runciman, W. G., Relative Deprivation and Social Justice (Berkeley, California: The University of California Press, 1966).
- _____, "'Social' Equality," The Philosophical Quarterly, 17(1967), 221-30.
- _____ and Sen, A. K., "Games, Justice, and the General Will," Mind, 74(1965), 554-62.
- Samuels, P. L., "An Examination of Rawls' Theory of Justice," Unpublished Ph. D. Dissertation (University of Rochester, 1972).
- Schaar, John, "Reflections on Rawls' Theory of Justice," Social Theory and Practice, 3(1974), 75-100.
- Scanlon, Thomas M. Jr., "Rawls' Theory of Justice," The University of Pennsylvania Law Review, 121(1972-73), 1020-69. A revised version is in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 169-205.
- Schaefer, David L., "The 'Sense' and Non-sense of Justice: An Examination of John Rawls's A Theory of Justice," The Political Science Reviewer, 3(1973), 1-41.
- _____, "A Critique of Rawls' Contract Doctrine," The Review of Metaphysics, 28(1974), 89-115.
- Schuler, G. F., "Rawls on Promising," Proceedings of the New Mexico-West Texas Philosophical Society, (April, 1974), 10-14.

- Schwartz, Adina, "Moral Neutrality and Primary Goods," Ethics, 83(1973), 294-307.
- Sen, A. K., Collective Choice and Social Welfare (San Francisco: Holden-Day Inc., 1970).
- _____, "Rawls versus Bentham: An Axiomatic Examination of the Pure Distribution Problem," Theory and Decision, 4(1974), 301-9. A slightly revised version is in Reading Rawls: Critical Studies of A Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), 283-91.
- Shue, Henry, "The Current Fashions: Trickle-Downs by Arrow and Close-Knits by Rawls," The Journal of Philosophy, 71(June 13, 1974), 319-27.
- _____, "Liberty and Self-Respect," Ethics, 85 (1975), 195-203.
- _____, "Justice, Rationality, and Desire: On the Logical Structure of Justice as Fairness," The Southern Journal of Philosophy, 13(1975), 89-97.
- Singer, Peter, "Sidgwick and Reflective Equilibrium," The Monist, 58(1974), 490-517.
- _____, Democracy and Disobedience (London: Oxford University Press, 1974), esp. 86-92.
- Slote, Michael, "Desert, Consent, and Justice," Philosophy and Public Affairs, 2(Summer, 1973), 323-47.
- Smith, James W., "Justice and Democracy," The Monist, 55 (1971), 121-33.
- Snare, Frank, "John Rawls and the Methods of Ethics," Philosophy and Phenomenological Research, 36(1975), 100-12.
- Soble, Alan, "Rawls on Self-Respect," an abstract, The Journal of Philosophy, 72(October 23, 1975), 629.
- Speigelberg, Herbert, "Good Fortune Obligates: Albert Schweitzer's Second Ethical Principle," Ethics, 85(1975), 227-34.
- Sterba, James, "Justice as Desert," Social Theory and Practice, 3(1974), 101-16.

- Sullivan, Daniel, "Rules, Fairness, and Formal Justice," Ethics, 85(1975), 322-31.
- Tapp, June L. and Levine, Felice, "Legal Socialization Strategies for an Ethical Legality," Stanford Law Review, 27(1974), 1-72.
- Tattershall, Gerald, "A Rawls Bibliography," Social Theory and Practice, 3(1974), 123-7. (12 entries by Rawls, 59 about Rawls.)
- Taylor, Paul W., "Universalizability and Justice," in Ethics and Social Justice, ed. by Howard E. Kiefer and Milton K. Munitz (Albany, New York: State University of New York Press, 1968), 142-63.
- Taylor, Richard, "Justice and the Common Good," in Law and Philosophy, ed. by Sidney Hook (New York: New York University Press, 1964), 86-97.
- Taylor, Telford, "The Concept of Justice and the Laws of War," The Columbia Journal of Transnational Law, 13 (1974), 189-207.
- Teitelman, Michael, "The Limits of Individualism," in "Symposium: A Theory of Justice," The Journal of Philosophy, 69(October 5, 1972), 545-56.
- Thibaut, John, Walker, Laurens, LaTour, Stephen, and Houlden, Pauline, "Procedural Justice as Fairness," Stanford Law Review, 26(1974), 1271-89.
- Tullock, Gordon, "Comment on Rae's 'The Limits of Consensual Decision'," The American Political Science Review, 69(No. 4, December, 1975), 1295-7.
- Van Dyke, Vernon, "Justice as Fairness: For Groups?" The American Political Science Review, 69(1975), 607-14.
- Wellbank, J. H., "Utopia and the Constraints of Justice," in Utopia/Dystopia? ed. by P. E. Richter (Cambridge, Massachusetts, Schenkman Publishing Co., 1975), 31-41.
- Wittman, Donald, "Punishment as Retribution," Theory and Decision, 4(1974), 209-38.
- Wolff, Robert Paul, "A Refutation of Rawls' Theorem on Justice," The Journal of Philosophy, 63(March 31, 1966), 179-90.

_____, Understanding Rawls, (Princeton University Press, 1977).

Zaitchik, Alan, "Hobbes and Hypothetical Consent," Political Studies, 23(December, 1975), 475-85.

_____, "Just Enough," Philosophical Quarterly, 25(October, 1975), 340-45.

APPENDIX

Arrow, Kenneth J., "Gifts and Exchanges," in Altruism, Morality, and Economic Theory, ed. by E. S. Phelps (New York: Russell Sage Foundation, 1975), 13-28.

Baumrin, Bernard H., "Autonomy in Rawls and Kant," Midwest Studies in Philosophy, (1976), 55-7.

Fullinwider, Robert K., "Hare on Rawls: A Worry About Possible Persons," Philosophical Studies, 29(1976), 199-205.

Hanfling, Oswald, "Promises, Games and Institutions," Proceedings of the Aristotelian Society, 75(1974-75), 13-31.

Harsanyi, John C., "Nonlinear Social Welfare Functions," Theory and Decision, 6(1975), 311-32.

Held, Virginia, "On Rawls and Self-Interest," Midwest Studies in Philosophy, Volume I: Studies in the History of Philosophy (1976), 57-60.

Kalin, Jesse, "Grice's Contract Ground and Moral Obligation: The Inadequacy of Contractualism," Philosophical Studies, 29(1976), 115-28.

Keenan, Francis W., "Justice and Sport," Journal of the Philosophy of Sport, 11(Spring, 1975), 111-23.

Luegenbehl, Dale, "Some Remarks on the Difference Principle," The Personalist, 57(1976).

Mason, H. E., "On the Kantian Interpretation of Rawls' Theory," Midwest Studies in Philosophy, Volume I: Studies in the History of Philosophy, Section IV, (1976), 47-55.

McClennen, Edward, "Comment" in Altruism, Morality, and Economic Theory, ed. by E. S. Phelps (New York: Russell Sage Foundation, 1975), 133-9.

Nagel, Thomas, "Comment" in Altruism, Morality, and Economic Theory, ed. by E. S. Phelps (New York: Russell Sage Foundation, 1975), 63-7.

Sterba, James, "Prescriptivism and Fairness," Philosophical Studies, 29(1976), 141-8.

Strasnick, Steven, "Social Choice and the Derivation of Rawls's Difference Principle," The Journal of Philosophy, 73(February 26, 1976), 85-99.

Urmson, J. O., "A Defense of Intuitionism," Proceedings of the Aristotelian Society, 75(1974-75), 111-19.

ABSTRACT

"THE 'GOOD' IN RAWLS' CONTRACTARIAN THEORY OF JUSTICE"

The aim of this study is to investigate Rawls' theory of moral justification by placing his concept of "Goodness as Rationality" in the context of his overall account of justice and morality. More specifically, my purpose is to determine how far Rawls' account has been able to undermine the challenge of moral scepticism. Does Rawls' account give an effective reply to the question "why be moral?" ? Has he succeeded in making good his claim that he has constructed a theory of justice that while remaining normative, is both descriptive and objective? Has he been able to show that moral judgments are among the kinds of judgments whose truth can be determined publicly by any rational observer who is acquainted with the facts?

CONCLUSIONS:

1. Rawls' attempt to revise the traditional contract theory does not succeed. He is unable to provide a moral basis for the social contract doctrine because his interpretation of the original position cannot yield principles that are uniquely acceptable from the moral point of view.
2. Rawls' method of reflective equilibrium does not succeed in the context of a prescriptive justificatory theory of justice. The method is only effective to the extent to which Rawls' theory can be regarded as a purely descriptive-explanatory theory of justice.

3. The two principles of justice as fairness need not commend themselves to all reflective persons as correct and impartial principles of justice. In particular the principles are influenced by certain ideologies of contemporary liberalism.
4. Rawls' account of the 'good' suffers from an inadequate theoretical basis. His attempt to extend his thin theory to the full theory of the 'good' does not work for at least one important reason: an effective sense of justice (one of the fundamental virtues which rational members of a well-ordered society would desire in one another) cannot be derived from Rawls' original position. In view of the motivational assumptions in the original position, as well as Rawls' reliance on an economic theory of rationality, an effective distinction cannot be made between a morally good person and a person of good morals.

One of the important lessons from the failure of Rawls' theory is that contemporary writers in ethics would be reminded once more about what is possible and what is not possible in ethics. The study also stresses the need for further investigation of the notion of rationality, in particular, how this concept applies to morality.