



National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

The Limits of Contractarianism: Rawls' Moral
Methodology and Ideological Framework

by Yuan-Kang Shih

Thesis Presented to the School of Graduate Studies
of the University of Ottawa as Partial Fulfillment
of the Requirements for the Degree of Doctor of
Philosophy

Ottawa, Canada, February 1980

Acknowledgements

I wish to express my deep gratitude to my thesis adviser, Professor Kai Nielsen. His most helpful comments and unflinching encouragements have made this thesis much better than it would have been. I also wish to thank Professors Hilliard Aronovitch, Christopher Morris, and Naomi Scheman for their useful comments on the draft of the thesis. I am especially indebted to Professor Morris, who read the whole draft and discussed with me on every issue in detail.

Contents

Introduction	1
Part I Rawls' Moral Methodology	
Chapter 1 Traditional Contractarianism Versus Rawlsian Contractarianism	16
I The Traditional Contractarianism	19
II The Principle of Fidelity	29
III The Rawlsian Contractarianism	36
Chapter 2 Is the Contractarian Method Valid for Doing Moral Philosophy?	61
I Social Contract, the Institution of Promising and the Obligation to Keep One's Promise	61
II A Dilemma	76
Chapter 3 The Justificatory Use of the Social Contract; Social Contract as the Justification of Moral Principles	85
I Social Contract and the Justification of Moral Theories	85
II Rational Choice and Moral Principles	93
III The Role of Justice	108
IV The Original Position and the Moral Point of View	113
Chapter 4 The Method of Reflective Equilibrium	131
I What is the Method of Reflective Equilibrium	131
II The Method of Reflective Equilibrium and Relativism	143
III The Method of Reflective Equilibrium and Intuitionism	162

Part II	Rawls' Ideological Framework	
Chapter 5	Contractarianism and Individualism	179
I.	An Elucidation of the Terms "Ideology" and "Individualism"	179
II	The State of Nature and Individualism	191
III	The Theory of Primary Goods	205
IV	The Circumstances of Justice and Individual- ism	218
Chapter 6	The Difference Principle and the Differences of Life-Prospect	238
I	The Derivation of the Difference Principle	238
II	The Maximin Rule and the Derivation of the Egalitarian Principle	255
III	The Difference Principle and Class-division	270
IV	Pure Procedural Justice and Fairness	280
Concluding Remarks		293
Bibliography		299
Appendix (Abstract)		307

Introduction

The history of modern philosophy is highlighted by rebellions against the practices of previous philosophers and by efforts to transform philosophy into a science, a discipline wherein there are universally accepted decision-procedures for testing philosophical theses. In Descartes, in Kant, in Husserl, and in Wittgenstein, we find the same claim being made that the philosophical theses presented are no longer opinions only, but have been elevated to the status of rigorous scientific knowledge. The way to achieve this, according to them, always consists in the adoption of a new philosophical method, for instance, Descartes' utilization of the methods of "clear and distinct ideas" and "universal doubt," Kant's "transcendental method," Husserl's "phenomenological bracketing," and the early Wittgenstein's attempts to dissolve philosophical problems by laying bare the logical forms of our languages. The aim of all of these revolutionaries is to replace opinions with knowledge in philosophical inquiry. The first and most important move in realizing this goal is to adopt

a new philosophical method which, in their opinion, will enable them to get rid of all the presuppositions embedded in previous philosophical theses.

Unfortunately, all these methodological innovations themselves have been found to be dependent upon certain substantive and controversial philosophical theses. In other words, the neutrality of the newly adopted method has never been successfully established. A revolutionary will adopt a new philosophical method only if he has already committed himself beforehand to certain substantive philosophical thesis. Everyone of them tries to be presuppositionless, but none seems to have succeeded. This piece of historical evidence leads us to the question of whether philosophical method is an integral part of a substantive philosophical thesis, so that the adoption of a certain method by a philosopher automatically commits him to the acceptance of a certain substantive philosophical thesis; and only the proponents of this philosophical thesis are willing to use the method in question as his guide in deliberating on philosophical problems. Historically speaking, the answer seems to be yes. However, our interest here extends beyond historical facts. We want

to know whether this is logically the case, i.e., whether the adoption of a philosophical method logically requires one's prior acceptance of a substantive philosophical thesis. Again, the answer seems to be yes. This is not really surprising in view of the nature of philosophy and the way philosophers have been doing it.

Let us start with the second point, i.e., how philosophers have been doing philosophy. Generally speaking, philosophers concern themselves with the most fundamental problems about the nature of human knowledge, or about the universe, or about morality, and so on. Aside from this very broad characterization of what they have concerned themselves with, there is very little they would share in common. The lack of agreement among philosophers is notorious. In philosophical discussion or dispute, philosophers tend to accuse their opponents of being mistaken on those fundamental issues of principle, such as the method they employed in conducting philosophical inquiry, or their view of the significance of philosophy, or even their view of what philosophy itself is and what it is supposed to be doing. As a matter of fact, they have been constantly disagreeing on what are the central

issues of philosophy, what is the proper method for doing philosophy, and what constitutes a proper solution to a philosophical problem. Actually, the situation is even worse than this. It is by no means uncommon among philosophers to dispute about what philosophy itself is. This lack of agreement on the fundamental matters among philosophers can be described as, following Kuhn, "the pre-paradigm stage of development." Whether this situation is due to the nature of philosophy and will be surpassed in the future need not concern us here. But because of this situation it makes every philosopher start anew from the beginning as if he or she were the first person to inquire about the problems. Students of science today hardly read anything by Aristotle; but philosophy students still have to start with Plato as if there had been no progress in philosophy since Plato, and this might very well be true. Because of this lack of consensus among philosophers, and the alleged necessity to start from the beginning in any kind of probing philosophical work, it is natural that philosophers tend to question the validity of the approaches taken by their predecessors. But how is one to know what is the proper method for philosophical

inquiry before one knows what philosophy is? Indeed, before we can make a methodological innovation, we have to know what philosophical problems are, and what philosophy is supposed to be doing. Yet these problems themselves are philosophical problems. So, it follows from this that a substantive philosophical belief must be arrived at before a new method can be adopted. Furthermore, the acceptance of this substantive philosophical thesis as to what philosophy is cannot be justified by appealing to the method one is going to adopt since it should rather be the other way round, i.e., it should be the case that we accept the substantive thesis of what philosophy is before we can know what method is appropriate for doing it. It follows that if a revolutionary wants to defend the philosophical conclusions by appealing to the new method he uses, he is open to the charge of being circular.

Secondly, let us consider the problem of whether the adoption of a new philosophical method presupposes the acceptance of a certain substantive philosophical position from a different angle, namely, the nature of philosophy itself. This consideration is by no means unconnected with the sociological considerations we have just

discussed. Philosophy is a highly self-reflecting discipline. By "self-reflecting discipline" we mean a discipline which takes itself as the object of its inquiry, that is to say, it is directed to the considerations of the questions like the following: "What the discipline itself is?" and "What is the significance of it?" and so on. Anyone who engages in philosophical speculation, sooner or later, will encounter such fundamental questions as "What philosophy is?", "What is the significance or value of philosophy?", "What constitutes a solution to philosophical problems?", and "What is the proper method for philosophical inquiry?", and so on. No matter how different their views might be on these issues, philosophers will not quarrel with the claim that philosophy is a highly self-reflecting discipline in the sense that they themselves are constantly engaged in speculating about these questions. These problems are proper philosophical problems. Anyone deliberating on them is engaged in philosophizing. As a matter of fact, it seems to be a necessity for a person who is serious about philosophy to be clear about these issues, and hopefully be able to come up with some answers to them. In contrast, science

is not self-reflecting. Scientists qua scientists do not have to answer questions like "What science is?", "What is the significance of scientific research?" and "What constitutes a solution to a scientific problem?", and so on. Since scientists qua scientists do not ask this type of questions, so science does not take itself as the object of its inquiry; rather, its research is directed toward something other than itself. For example, astronomy is the discipline which investigates the movement of heavenly bodies. An astronomer does not have to ask questions about astronomy itself, and as a matter of fact, he does not. Of course, during a period of scientific revolution, scientists do address themselves to such fundamental questions about science itself. But in this case they cease to act simply as scientists. Rather, they enter the territory of philosophy. Those fundamental questions are not scientific questions proper. Scientists qua scientists normally just do not raise them. When they do, they are no longer engaged in scientific research. On the other hand, those fundamental questions about philosophy are themselves philosophical questions, and anyone who raises them is still engaged in philosophizing.

In the light of this analysis, we can get back to our original problem, namely, does a methodological innovation in philosophy presuppose the prior acceptance of a certain substantive philosophical position? And if so, why? Since philosophy is highly self-reflecting in the sense specified above, in philosophizing we constantly address ourselves to the set of fundamental questions about philosophy itself. Among this set of questions, the most fundamental one is "What is philosophy?" The rest of them in certain important respects, depend on whether we can come up with an answer to it and what kind of answer we come up with. Indeed, before an answer to this question is arrived at, i.e., before we can say what philosophy is and what it is supposed to be doing, there seems to be no way that we can give answers to other questions in the set. It is logically impossible for us to ask questions like, let alone answer them, "What is the proper method for conducting philosophical inquiry?", and so on, unless we have an answer to the most fundamental question "What is philosophy?" However, this most fundamental question itself is a philosophical question. Philo-

sophers who ponder on it are engaged in philosophizing and the answer to it is a substantive philosophical thesis. So, we may conclude from this analysis that philosophical methodology is dependent on a substantive philosophical thesis or set of theses in the sense that anyone who uses a certain philosophical method must already have been committed to a certain substantive philosophical position. This conclusion is not really surprising in view of the nature of philosophy. That is to say, it is highly self-reflecting and hence takes itself as the object of inquiry.

Recently, Rawls tries to revive the contractarian approach to moral philosophy. His aim, like all the previous revolutionaries in philosophy, is to transform this branch of philosophy into a rigorous science, and to replace opinions with knowledge. Thus he says:

The argument aims eventually to be strictly deductive. To be sure, the persons in the original position have a certain psychology, since various assumptions are made about their beliefs and interests.....But clearly arguments from such premises can be fully deductive, as theories in politics and economics attest. We should strive for a kind of moral geometry with all the rigor which this name connotes.

Nevertheless, Rawls differs in one crucial respect from the giant figures in the history of philosophy. While

they all strived to be presuppositionless, Rawls, in deriving the principles of justice, boldly utilizes numerous premises drawn from other disciplines as long as these premises look "reasonable," "natural" and "intuitively acceptable" to him, and hopefully to the readers as well.

In his own words,

Moral philosophy must be free to use contingent assumptions and general facts as it pleases. There is no other way to give an account of our considered judgments in reflective equilibrium. This is the conception of the subject adopted by most classical British writers through Sidgwick. I see no reason to depart from it.²

There are two methods Rawls employs in Constructing and justifying his theory of justice.³ First, there is the contractarian method. It is the use of this method which enables Rawls to make the claim that his theory is in line with the tradition of the social contract theory. Second, there is the method of reflective equilibrium. The employment of this method gives Rawls' theory a heavy Quinean tone.

The contractarian method uses the notion of a social contract as the foundation of a moral theory. Rawls argues that moral principles are to be derived and justified by

a social contract among rational people. The method requires us to imagine that a certain group of people coming together to make a social contract. Rawls calls this starting point, which corresponds to the state of nature in traditional social contract theories, "the original position." There are certain conditions imposed on the original position which Rawls thinks would be fair to each and every contractor as free and equal participant (hence, the name "Justice as Fairness"). The conditions imposed on the original position can be divided into two parts. First, on the subjective side, the contractors are rational, self-interested, and temporarily behind a "veil of ignorance," which makes them have access only to theoretical knowledge. The veil prevents them from knowing who they are, what their talents are, and what their tastes, ambitions, and religious beliefs are, and so on. Secondly, on the objective side, there is the condition of relative scarcity of resources so that cooperation is both necessary and possible. Rawls argues that the principles unanimously chosen by the contractors in the original position will be principles of justice.

The second method Rawls utilizes is the method of reflective equilibrium. The method states that moral

principles are justified if they can be shown to match our lower-level "intuitive" judgments about particular cases. These lower-level judgments are supposed to be derivable from the principles in question. In using this method to justify moral principles, there are certain conditions that must be observed. First, not all lower-level judgments are to be used -- only those we make with great confidence concerning our basic moral convictions. They should be based on true beliefs, and untainted by disturbing factors such as our special interests, and so on. Second, alternative accounts of such judgments must be taken into consideration. Rawls refers to the state in which the principles match the "considered judgments" as the "reflective equilibrium." It is reflective since the alternatives have been considered; and it is an equilibrium because the principles and judgments coincide.

The outcome of Rawls deliberation is a highly original normative ethical theory. Rawls argues that two principles will be agreed upon by the contractors in the original position to regulate the basic structure of society. They are:

- (1) 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.

- (2) Social and economic inequalities are to be arranged so that they are both (a) to the greatest benefit of the least advantaged, consistent with the just savings principle and (b) attached to offices and positions open to all under conditions of fair equality of opportunity.

The aim of this dissertation is to provide a critical examination of Rawls' theory of justice from two perspectives. First, we shall carefully examine his moral methodology. It will be argued that neither of the methods will work as Rawls intends. We shall contend that Rawls' claim that moral principles are the objects of rational choice in the original position is problematic. The fact that the principles will be chosen by the contractors in the original position does not guarantee that they are moral principles. As to the method of reflective equilibrium, we shall argue that it does not really carry us beyond relativism and intuitionism. Secondly, we shall be concerned with the problem of Rawls' ideology. An attempt will be made to show that the conditions imposed on the original position embody certain prejudices. Moreover, in accordance with what we remarked earlier that moral methodology is part of substantive philosophical thesis, we shall argue that the methods themselves are not neutral in the sense that not every moral philosopher can and will

use them. Only those who are committed to certain substantive philosophical theses will employ these methods for doing moral philosophy.

Footnotes

1. John Rawls, A Theory of Justice (Cambridge, Massachusetts: Harvard University Press, 1971), p.121.
2. Ibid p.51.
3. Actually, there is a third method Rawls uses in consolidating his theory, namely, the method of congruence. It is roughly as follows: we have a conception of the good, and this conception of the good and our conception of justice are mutually supportive. Hence, Rawls argues, to act justly is in our own interest. But when it comes to the derivation and justification of the principles of justice, the first two methods play the decisive role.
4. Rawls, op. cit., p.60.

Part I Rawls' Moral Methodology

Chapter 1 Traditional Contractarianism versus Rawlsian Contractarianism

In A Theory of Justice, Rawls has not systematically presented his methodology. He says, "Occasionally there are methodological comments and asides, but for the most part I try to work out a substantive theory of justice."¹ This creates some difficulties for us in dealing with his moral methodology which is certainly one of the most important parts of his theory in the sense that, if the methods he uses in constructing the theory are themselves unsound, then the theory built will be greatly jeopardized. However, from his occasional remarks on the topic in the book and an earlier essay entitled "Outline of a Decision-Procedure for Ethics,"² which deals exclusively with moral methodology, we can draw a fairly clear and complete picture of his moral methodology. As we said earlier, there are two basic methods Rawls employs in constructing his theory of justice, namely, the contrac-

tararian method and the method of reflective equilibrium. In part I, we shall examine in detail what these methods are, what uses Rawls makes of them, and whether they really accomplish the tasks Rawls assigns them. We shall not be concerned with the problem of whether these two methods are independent of each other or are integral parts and inseparable from one another. For the purpose of analysis, we can treat them separately. First, we shall deal with the contractarian method. As to this method, several problems will be discussed extensively:

(1) What is a contractarian doctrine in the traditional sense of this term? Are there different sorts of contractarian doctrine? If so, what is the common denominator underlying all of them so that they all warrant this name?

(2) What is the Rawlsian version of contractarianism? Does it fit in with the traditional social contract theory so that it also warrants the name "contractarianism?" Do they share the same logic? We shall also discuss the meaning of Rawls' claim that the aim of his theory is "to present a conception of justice which generalizes and carries to a higher level of abstraction the familiar theory of

the social contract as found, say, in Locke, Rousseau, and Kant."³ Our interpretation of this claim is that the concept of social contract of which the traditional contract theorists made use in political philosophy is utilized by Rawls in coping with moral problems. In other words, the original contract, instead of being the foundation of political society, becomes the foundation of moral principles.

(3) Is contractarian method valid for doing moral philosophy? Does it make sense to say that moral principles are the objects of our choice?

(4) We shall examine Rawls' claim that the principles chosen by the people in the original position are principles of justice. This leads us to inquire into the device of the original position and the various conditions imposed on it.

We shall in the chapters following our discussion of the contractarian method examine the method of reflective equilibrium. With respect to this method, the following problems will be discussed:

(1) What is the equilibrium method?

(2) Rawls claims that his moral theory is universal-

istic, i.e., it is acceptable to people from different cultural backgrounds. But is it really so? Does the method of reflective equilibrium enable us to advance beyond moral relativism?

(3) The method of reflective equilibrium and intuitionism.

I. The Traditional Contractarianism

In his book In Defense of Anarchism, Robert Paul Wolff says: "the discovery, analysis, and demonstration of the forms and principles of legitimate authority -- of the right to rule -- is called political philosophy." The traditional social contract theorists would not quarrel with Wolff over this statement. They viewed their fundamental tasks as the establishment of principles on which the legitimacy of political authorities rest and the search for a valid foundation for political obligations which individuals are supposed to have under certain circumstances. Actually, these two problems are just two sides of the same coin. When a political institution (government, state, etc.) has certain legitimate authority over an individual,

the individual will correspondingly have certain obligations to it. When confronted with these problems, the traditional social contract theorists had a clear and unequivocal, if unconvincing, answer:

A political institution (state) S has certain authority over a person P under circumstance C if and only if P has entered into a social contract A and the fulfillment of A implies that under C, S has the said authority over P.⁵

and correspondingly,

A person P has a political obligation to perform an action N under circumstance C if and only if P has entered into a social contract A and the fulfillment of A implies that under C, P is required to perform N.⁶

According to these formulations, the legitimacy of a political authority and a person's political obligations are based on a social contract. Because we have entered into a certain contract, and the fulfillment of that contract implies both that a certain political institution has authority over us and that we have an obligation to perform or refrain from performing certain action under certain circumstances, therefore, the authority is legitimate and our obligations are justified. To put it briefly, political authorities and obligations are explained and justified by the existence of a social contract we have

entered into. Since a contract is a sort of agreement, and in general, we have an obligation to keep our agreement, consequently, we acquire a political obligation to perform certain actions under certain circumstances through an obligation we generally acknowledge, i.e., we should keep our agreements. So, according to this theory, political obligations are based upon our personal obligation to keep our agreements.

However, a social contract theory, formulated in this manner, can easily be disposed of.. To dismiss the theory, we can simply point out the lack of historical and factual evidence that any one (or most) of us has entered into any such social contract. Hence, the alleged political obligations we have or the authority the state has over us, are not, given that account, really warranted.

Most traditional contract theorists who realized this difficulty in their theories, and were somehow still convinced that the notion of a social contract was useful both in explaining and justifying political authority and obligation, resorted to the concept of a tacit consent. By positing this additional element, they would still have a full-blown social contract theory. This new premise

states that individuals living in a society usually have tacitly given their consent to enter into a social contract even if an overt agreement is not there. The tacit consent implies that members of a society have entered into a social contract. It is as good as an overt contract. Hence, the political authorities and obligations are, again, based on a contract. Thus, when Socrates told Crito that "But he who has experience of the manner in which we order justice and administer the State, and still remains, has entered into an implied contract that he will do as we commend him,"⁷ he was clearly laying out this version of the social contract theory. According to this version, .

A political institution (state) S has certain authority over a person P under circumstance C if and only if P has performed certain action N, and P's performance of N counts as a tacit consent of a social contract A by P, and the fulfillment of A implies that under C, S has the said authority over P.

and correspondingly,

A person P has a political obligation to perform an action N under certain circumstance C if and only if P has performed certain action N', and P's performance of N' counts as a tacit consent of a social contract A by P, and the fulfillment of A implies that under C, P is required to perform N.

Just what counts as a tacit consent by a person to enter into a social contract is a matter of controversy. The passage we quoted above from Plato states that a person has given his tacit consent to enter into a social contract with other people when he has experience of the manner by which the group of the people order justice and administer the State, and still remains in the State. The consent usually consists of the acceptance of benefits and advantages from an institution or a law. Since the acceptance of benefits and advantages constitutes our tacit consent to enter into a social contract, therefore, if we have received benefits from an institution, and if the law says that the institution has certain authority over us, or, that we are obligated to perform or refrain from performing certain action under circumstance C, then it should be so even though we have not overtly entered into any social contract. Again, since a contract is an agreement, and in general, we have an obligation to keep our agreements, consequently, we acquire a political obligation through a personal or moral obligation we generally acknowledge, i.e., we should keep our agreements. Although this version of the social contract theory needs an additional step to

justify the legitimacy of political authorities and obligations, it ultimately advocates the thesis that they are based on a social contract.

Having formulated clearly the two versions of the traditional social contract theory, the question now is: When can we label a theory in political philosophy as a social contract theory? What are the characteristics of this type of theory? What is the common denominator which underlies these two versions of the traditional social contract theory? Is there any other type of social contract theory? What we need now, if such can be had, is a formulation which states the necessary and sufficient conditions for a social contract theory. Let us try some. If we maintain that for a theory to qualify as a social contract theory, it must assert explicitly that an actual contract exists, then it will obviously rule out the tacit consent theory as social contract theory. Nevertheless, it is commonly held that a social contract theory does not commit the proponents to the existence of an actual social contract. The condition is definitely too strong to be a necessary one. Now, suppose, instead of stating that an actual contract must have existed, we weaken the condition by

stating that only a tacit contract is required, will this condition still be too strong? Some people will be apt to think so, since, according to this formulation, it will still rule out some theories which they would like to consider as contractarian, e.g., the hypothetical contract theory.⁸ Our intuition here seems to be that for a theory to be contractarian, it does not have to assert the existence of a social contract, either actual or tacit, as long as the notion of a social contract plays some role in the theory. Then, will it be sufficient for a theory to be contractarian as long as a reference is made to a social contract? That is to say, the words "social contract" appear somewhere in the theory? This formulation, though satisfactory with respect to the inclusion of all three versions of contractarian doctrine, runs into its own difficulties. It is obviously too weak to be sufficient. If this formulation is adopted, then any kind of theory about political obligations and authorities will count as contractarian. For instance, we will have to take the following perfectionist theory as contractarian if we accept the formulation:

A political institution (state) S has certain authority over a person P under circumstance C if and only if S's having the said authority

over P under C will maximize the achievements of human excellence in art, science, and culture, whether or not the said authority follows from a social contract.

correspondingly,

A person P has a political obligation to perform an action N under circumstance C if and only if P's performance of N will maximize the achievements of human excellence in art, science, and culture, whether or not the requirement of the performance of N by P follows from a social contract.

Classical utilitarianism can also be conceived as contractarian if we substitute "total utility" for "achievements of human excellence in art, science, and culture" in the above formulation of perfectionism. If any theory can count as contractarian, the label itself becomes useless. We no longer can distinguish contractarian doctrines from non-contractarian ones; yet, intuitively, we do think there must be a difference between them. In characterizing a certain type theory as contractarian, what we need to do is to find out the essential features of this type of theory which will differentiate it from other types of theory. In the first place, we have some vague idea as to what a contractarian doctrine is, the task for us now is to sharpen the idea so that a clear and distinct form-

ulation can be arrived at. Now let us proceed on this line.

We stated earlier, following Wolff, that the principal task of political philosophy consists in "the discovery, analysis, and demonstration of the forms and principles of legitimate authority." We also stated that the traditional social contract theorists considered their theories as solutions to this problem and the parallel problem of how to justify political obligations. By positing a "social contract," either actual or tacit, or even hypothetical in some cases, according to them, political authority and obligation are both explained and justified. The notion of a "social contract" is the basis of authority and obligation in the sense that they are derived and originated from it. So, in characterizing social contract theory, we should not neglect this feature of the theory. The formulation must always capture the idea of the indispensability of the contract in explaining and deriving of the political authorities and obligations. It follows that in defining the social contract theory, the mere appearance of the words "social contract" in a theory will not be sufficient for that theory to qualify as contractarian.

For a theory to be contractarian, the concept of social contract must be doing some work in the theory in deriving political obligations instead of simply sitting there idly and doing nothing. Our formulation of the perfectionist theory above shows the idleness of the notion of social contract in that theory. Since social contract theories are principally designed for the derivation and justification of political authorities and obligations, the work the notion of a social contract is supposed to do must be indispensable in their derivation and justification. Accordingly,

A theory in political philosophy is contractarian if and only if it states that the notion of a social contract in that theory is both a necessary and a sufficient condition for the derivation and justification of political authorities and obligations.

Defined in this manner, as we can see, both versions of the traditional contract theory, i.e., the overt consent⁹ and the tacit consent theory, qualify as contractarian. Since, according to them, political authorities and obligations are derived solely from a social contract, to say that Y is derived from X is tantamount to saying that if we have X, we will have Y. Hence, if there is a contract,

10

there will be political authorities and obligations.

Furthermore, since only from a social contract can they be derived and established, therefore, a social contract is a necessary condition for them. That is to say, if there is no social contract, there will be no political authorities and obligations. However, the perfectionist doctrine does not satisfy this formulation, since, in that doctrine, we can delete the reference to the notion of a social contract and yet still have all the political authorities and obligations with us. Hence, they must be based on something other than a social contract.

II. The Principle of Fidelity

Having made clear what we mean by "contractarianism," we now turn to the question of what the contract theory must presuppose in order to guarantee that the derived political authorities will be recognized and the obligations acknowledged? Can the contract itself alone guarantee them? Does the validity of the contractarian approach to the issues of political authorities and obligations rest on something other than a contract? We shall argue that it must pre-

suppose the principle of fidelity, i.e., for the validity of this approach to be insured, it must presuppose the validity of the principle that promises or agreements must be kept. Moreover, we will show that the validity of this principle must be based on something other than a social contract.

According to the contractarian theories, when a certain number of people decide to make a social contract, which is to establish the authorities of political institution and assign rights and obligations to each and every contractor, then the authorities, rights, and obligations established by the contract will come into being and they will be justified as well. But we are puzzled by the question "How can a contract create authorities and obligations where none existed before?" Does the contract create and justify them by itself alone? Is it self-contradictory to ask the following type of Moorish question: "According to the contract I entered, it is my obligation to perform A, but do I really have an obligation to perform A?" Most of us tend to think that this question, though not self-contradictory, is a little out of line. If a contract we entered into says that we should perform an action A, then

it seems to us, we do have an obligation to perform A. But is this really so? Can we assert this statement without presupposing something else? For instance, the validity of the principle of fidelity. Let us consider an example. Suppose A, B, and C live in the same house. They agree that they will take turns to clean the house. Each person is to be responsible for two days, and for Sunday, everybody takes a day off. A's turn is Monday and Tuesday, B cleans on Wednesday and Thursday, and C is responsible for Friday and Saturday. Now, it is the first week of their cooperation. A and B did their jobs on the previous days in accordance with the agreement. Today is Friday. A and B plan to entertain some friends on that night. When they come home after work, they find that the floor is filthy, and the garbage piled up, and C is sitting there puffing his pipe and listening to music. With a little anger, they ask C why he has not cleaned the house. Much to their surprise, C replies, "Why should I clean the house?" Now they become really angry, and shout at C: "You agreed to the terms that each of us takes turn to clean the house, and it is your turn today." But C replies with calm, saying that, even if he made an agreement with them to clean the house

on Friday and Saturday, it does not create an obligation for him to do so because in the agreement, nowhere it says that he should keep his agreement. Therefore, he does not feel obliged to clean the house. A and B answer him furiously, "But an agreement is meant to be kept, so go ahead and clean, etc." In this fictional case, we find C evasive and lacking in character because, intuitively, we do agree with A and B, i.e., we think that usually if we make a promise or enter into an agreement, we should be doing so with the intention of keeping it, otherwise there will be no point in making a promise or entering into an agreement.¹¹ Indeed, some philosophers even find it self-contradictory to ask the question: "How can a promise create obligations?" John Searle, for example, maintains that promising is an institution one of whose rules requires that it be kept. To ask how a promise can create obligations is like asking¹² how a touchdown can create six points in football game. But when philosophers like Hume and Prichard ask how promises can create obligations, we do not find them evasive or lack of character; on the contrary, we find it a genuine puzzle concerning the problem of the creation of obligations. Contractarians believe that ~~all~~ our political obligations

are derived from the contract we entered. But if the contract itself does not state or presuppose that we are obligated to keep our agreement, then, what is the basis of the principle that we should always keep our agreement? And from where can it be derived? It cannot be based on another contract, since if we try to avoid the difficulty by setting up another contract which says that all our agreements must be kept, then the validity of this higher level contract must be sought in some other place. If we go on again in this manner, it is going to involve us in an infinite regress. Prichard's doubt about the usefulness of this kind of manoeuvre is precisely to the point when he says:

Promising requires the actual use of the word 'promise' or else of some equivalent such as 'undertake,' 'agree,' 'give you my word,' or 'will' in 'I will.' This being so, we can at least say that when I promise X to do some action, I am causing X to hear a certain noise, which has a definite meaning both to X and to me, together with the term ordinarily used for the action, in such a way that X believes that the sounds proceed from me. But then the question at once arises: "How can my doing this give rise to an obligation to do the action?" And the answer seems to have to be: "Only because I have already promised not to cause a noise of that kind in connection with the phrase for some action without going on to do the action." And if this answer

is right, what we call promising to do some action appears to be causing someone to hear in connection with the term for the action a noise of a kind which I have promised never to use in this kind of way without going on to do the action; and, again, the reason why, when I say to X "I promise to do so and so," I am bound to do the action is that I have previously made him a certain general promise.¹³

Thus, it seems, according to Prichard, the contractarian approach to the problem of the creation of obligations will bring us into a predicament. The invocation of a contract or a promise can give rise to an obligation only insofar as we appeal to a previous and more general agreement. But this manoeuvre of appealing to a previous agreement will not be able to rescue us from the difficulty, since we then have to seek the basis for the validity of this more general agreement. Therefore, contrary to what we intended, it will bring us into a process of infinite regress. This seems to lead us to think that if we want to build a solid basis for the authority of political institutions and our political obligations, we must go out of the contract and eventually appeal to something other than a social contract. And this is precisely what the traditional social contract theorists did. We stated earlier in connection with the formulations of the overt

and tacit consent theories that political authorities and obligations derived from a social contract, according to the traditional theorists, are eventually based on our personal obligation to keep our agreements. The principle itself, namely, that we have an obligation to keep our agreements, does not follow from the contract; rather, it is something for the validity of the social contract we must presuppose beforehand.¹⁴ Since this obligation itself is not part of the social contract, so it is not a political obligation; instead, it is a moral obligation which is supposed to be applicable to all kinds of agreements and promises. This is why some have argued that, at heart, social contract theorists believe that a political obligation is based on a moral obligation, namely, one must keep one's promises. This moral obligation, as we have pointed out, following Prichard, cannot be based on an agreement, or we will be brought into an infinite regress. We must stop somewhere in this process of going back in the search of foundations for our political obligations. The place we can stop is where we can find something which will be able to give us guarantee that all agreements and promises must be kept, and yet this thing itself cannot be a social contract.

Historically, the social contract theory ends up with the theory of natural law, i.e., it presupposes the natural law theory. Personal and moral obligations, according to this theory, are said to be part of, and to stem from, the natural law. Of course, the natural law itself is not something that is agreed to by a social contract. Its foundation lies somewhere else. The necessity for the social contract theory to appeal to something other than a social contract in order to guarantee the validity of the contract also explains the fact that the social theory of government and political obligations in this tradition is inextricably bound up with natural law theory.¹⁵

III. The Rawlsian Contractarianism

Now we turn to the questions of what is the Rawlsian version of contractarianism? Whether or not it fits in with the traditional social contract theory as Rawls claims it does? Does it share the same logic with the traditional contract theory? In view of the definition we worked out before for "social contract theory," we shall also ask the crucial question: Whether or not Rawls' theory

is contractarian? Our contention is that it is not, and that it does not share the same logic with the traditional social contract theory as far as the derivation of political authorities and obligations is concerned. Furthermore, we shall also discuss the meaning of Rawls' claim that his theory is to "present a conception of justice which generalizes and carries to a higher level of abstraction the familiar theory of the social contract as found, say, in Locke, Rousseau, and Kant."¹⁶

But before beginning our search for answers to these questions, we shall clear up a few things and try to keep them in mind firmly. Rawlsian contractarianism is different from the traditional social contract theory in two essential aspects. First, in contrast to the traditional social theory, it is a hypothetical contract theory; it does not assert the existence of a social contract in any form, either overt or tacit. Second, unlike the traditional social contract theory, it is used to derive moral principles rather than the principles of political authorities and obligations. In other words, it is a moral theory instead of a political theory. This latter difference between Rawls' theory and the traditional social contract

theory turns out to be crucial. We shall argue that the contractarian approach, while valid in dealing with the problems of political authorities and obligations, is not valid for doing moral philosophy. However, for the moment, we shall ignore this point, and go on to discuss the Rawlsian version of contractarianism as if the method could also be validly employed to deal with moral issues. At the end of this chapter, we shall come back to discuss this difference. In the next chapter, we shall question the validity of the contractarian approach for doing moral philosophy, and, there, a contrast between its use in deriving political authorities and obligations and in deriving moral principles will be made.

We said earlier that Rawls calls his own theory "a hypothetical contract theory." By this he means that his theory uses the notion of a social contract in deriving the principles of justice; yet it does not commit itself to the existence of an actual or tacit contract. In other words, the contract in question is a hypothetical one. Let us quote some passages from him:

The principles of justice...are the principles that free and rational persons concerned to further their own interests would accept in an initial position of equality as defining the fundamental terms of their association. 17

Our social situation is just if it is such that by this sequence of hypothetical agreements, we should have contracted into the general system of rules which define it. 18

We shall want to say that certain principles of justice are justified because they would be agreed to in an initial situation of equality. I have emphasized that this original position is purely hypothetical. 19

Since the original position is not conceived to be an actual historical state of affairs, the agreement reached by the contractors situated there cannot be an actual one either. The original position is hypothetical, therefore the agreement is hypothetical too. This is why Rawls calls his theory a "hypothetical contract theory."

As to the original position, he says:

In justice as fairness the original position of equality corresponds to the state of nature in the traditional theory of social contract. The original position is not, of course, thought of as an actual historical state of affairs, much less as a primitive condition of culture. It is understood as a purely hypothetical situation characterized so as to lead to a certain conception of justice. 20

There are, as I have said, many possible interpretations of the initial situation. This conception varies depending upon how the contracting parties are conceived, upon what their beliefs and interests are said to be, upon which alternatives are available to them, and so on. In this sense, there are many different contract theories. Justice

as fairness is but one of these... This most favored, or standard, interpretation I shall refer to as the original position. 21

Among the essential features of this situation is that no one knows his place in society, his class position or social status, nor does anyone know his fortune in the distribution of natural assets and abilities, his intelligence, strength, and the like. I shall even assume that the parties do not know their conceptions of the good or their special psychological propensities. The principles of justice are chosen behind a veil of ignorance. 22

These extensive quotations from Rawls' text should give us some idea of what the Rawlsian version of contractarianism is, even though it is definitely not sufficient to capture the totality of his powerful theory. Essentially, Rawls is asking us to imagine a hypothetical situation which is described as the original position, wherein a group of people are there to launch a joint adventure. They realize that subsequently they can all benefit from this cooperation. This joint adventure is the search for a set of principles which assign duties and rights to each and every contractor in the actual society. Rawls claims that they will arrive at a set of principles and they are the principles of justice. Obviously, the contract as conceived by Rawls is neither an actual contract nor a

tacit one since people have not really been in the original position, nor have they made the agreement. It is merely a hypothetical contract that people engaged in this "hypothetical thought experiment" would agree to if certain initial conditions were properly stipulated. Of course, as Rawls says, there can be different stipulations of this initial situation (the state of nature). The original position is but one of them. So, in contrast to the traditional social contract theory, which asserts the existence of a social contract reached by the contracting parties, either by overt or by tacit consent, the Rawlsian contractarianism does not assert the existence of a contract in any manner. All it says is that a contract would be agreed to if the initial situation were defined properly.

From this we may say that for Rawls:

A political institution (state) S has certain authority over a person P under circumstance C if and only if there is a hypothetical social contract A, and if P were in the Rawlsian original position, P would enter into A, and the fulfillment of A implies that under C, S has the said authority over P.

and correspondingly,

A person P has a political obligation to perform an action N under circumstance C if and only if there is a hypothetical social contract A, and if P were in the Rawlsian

original position, P would enter into A, and the fulfillment of A implies that under C, P is required to perform N.

Now we want to examine the logic of Rawls' theory concerning its derivation of political authorities and obligations and find out whether or not it fits in with the traditional theory of social contract, i.e., whether Rawls' theory qualifies as a contractarian doctrine according to the definition we worked out earlier for the term. We will do so by revealing the genuine basis of political authorities and obligations as they are derived in his theory. The definition states that a theory is contractarian if and only if the notion of a social contract is both a necessary and sufficient condition for the existence of political obligations and authorities, i.e., a theory is contractarian if and only if it asserts that a social contract and only a social contract can create political authorities and obligations. One thing stands out here is that it does not specify what kind of social contract it has to be. So, presumably, it can be either overt, or tacit, or even hypothetical. As long as the notion of a social contract is indispensable to the theory, it will be sufficient. As we argued before, both versions (the

overt and the tacit contract theories) of the traditional theory of social contract satisfy this condition. What about the Rawlsian contractarianism? We had some doubt then, and we shall elaborate on it now.

When a traditional social contract theorist is asked by someone: "Why am I obliged to perform an act A?", he always has the ready answer: "Because you have entered into a contract (either overtly or tacitly), and the contract says that it is your obligation to perform A." Of course, the validity of the principle of fidelity is taken for granted by both parties. In this way, the question is settled as long as there exists such a contract. But for a hypothetical contract theorist, the availability of the contract is lacking. When confronted with the same question, he cannot simply point to the contract and say what the traditional social contract theorist said, since no such contract exists. As Ronald Dworkin says: "A hypothetical contract is not simply a pale form of an actual contract, it is no contract at all."²³ The argument the hypothetical theorist can offer will be as follows: you would agree to a social contract if such and such conditions are satisfied and the hypothetical contract says that you should

perform A, therefore, it is your obligation to perform A. But the person will still be puzzled as to why a contract (hypothetical) which he has not entered into will have any binding force on him. It is just a little odd for us to think that we are obligated to do something if there is a hypothetical agreement which says so, even though we have not made the agreement. Let us consider an example. Suppose A drove his car to Toronto and when he came back from the trip, he asked B to pay half of the expenditure on gas. Stunned by this request, B naturally replied, "Are you crazy? Why should I?" A said to him, "Because if you had gone with me, you would have agreed to pay half of the gas bill, and since this is a hypothetical agreement, therefore it is your obligation to pay me." In this example, we find A's claim on the obligation B allegedly has sounds not only unjustified but also absurd. Intuitively, we just do not go along with A even if B would agree to pay half of the gas bill had he gone to Toronto with A. We do not find B evasive or dishonest in refusing A's demand. This amounts to saying that, in this case, the so-called hypothetical agreement does not create an obligation for B. When a hypothetical contract theorist says that a person

has an obligation to perform an act A, he argues as follows. If certain conditions $C_1, \dots, C_n$ are satisfied, then a person would agree to perform an action A. But in the above imaginary case, B can agree with this whole-heartedly and still refuse to pay the bill without contradicting himself by simply claiming that but $C_1, \dots, C_n$ are not met. However, sometimes a hypothetical agreement does seem to have binding force. If one does not do what he or she is supposed to do according to the hypothetical agreement, he or she will indeed sound evasive or dishonest. We can imagine a small boy asking his big brother to help him with some mathematical homework, and the big boy refuses because he wants to watch the hockey game. The small boy says, "If Mom were home, you would agree to help me." The big boy's reply -- "But Mom is not home" -- indeed sounds evasive. Here, the presence or absence of the mother is irrelevant as far as the duty to help a kid brother when he has difficulty is concerned. The reason the big boy's response seems evasive is because he is taking advantage of certain situation while this situation really should have no bearing on the duty he is supposed to have. So, when a hypothetical contract theorist like Rawls argues that a hypo-

thetical contract can "create" obligation even if the conditions $C_1, \dots, C_n$ which will lead to the agreement were not satisfied, it is clear that the alleged obligation is one of the obligations we already have. It is not a new obligation "created" out of a contract which has never been entered into, and will not be entered into. The reason it will never be entered into is because if it would be, then it is not a hypothetical contract any more; it becomes an actual contract. So, the way a hypothetical contract "creates" obligations is by pointing out, in an unusual way, the obligations we already have. It reminds us of the things we should do whether or not the contract will ever be entered into. It does not create new obligations. This is done by pointing out that, even if the conditions $C_1, \dots, C_n$ were never met (can never be met in Rawls' case) due to some intervening factors, but since they are the only things that are relevant when we consider the hypothetical agreement, so, by some artificial device, we try to isolate $C_1, \dots, C_n$ from the rest of the factors, and see what can be derived from them.

This turns out to be precisely the way Rawls sets up his theory. It is also the principal force of his

contractarianism. The principles the contractors in the original position are concerned with are the principles of justice. Though the conditions imposed on it are not satisfied in ordinary life, they are, according to Rawls, the only relevant factors as far as the choosing of the principles of justice is concerned. So, even if they are not satisfied in our ordinary world, they must be simulated in our "thought experiment" when we try to reach an agreement on what the principles of justice are. To put it more briefly, the whole "Rawlsian thought experiment" is just a dramatization of what we should acknowledge as the principles of justice in our ordinary life. So the aim of the device of the original position is,

...to rule out those principles that it would be rational to propose for acceptance, however little the chance of success, only if we knew certain things that are irrelevant from the standpoint of justice. 24

By contrast with social theory, the aim is to characterize this situation so that the principles that would be chosen, whatever they turn out to be, are acceptable from a moral point of view. 25

From these two passages, we can say that what Rawls has in mind when he characterizes the original choice situation is to rule out, in his view, the factors which are

irrelevant from a moral point of view. By doing this, the agreement (principles) reached will be guaranteed to be acceptable from the point of view of justice. This answers the question of why we should take any interest in the original position,

Finally, we may remind ourselves that the hypothetical nature of the original position invites the question: why should we take any interest in it, moral or otherwise? Recall the answer: the conditions embodied in the description of this situation are ones that we do in fact accept. Or if we do not, then we can be persuaded to do so by philosophical considerations of the sort occasionally introduced. 26

Consequently, even if the conditions imposed on the original position are not met in ordinary life, but since we accept them as the only relevant conditions from the point of view of justice, and will take an interest in the original position, it follows that the principles reached there will have binding force on us.

Now let us take a concrete example and see how this type of reasoning goes. Suppose a person fails to assist his neighbors when they are in need, even though it does not demand a drastic sacrifice on his part. Rawls will tell him that he fails to act in accordance with the principle of mutual assistance and it is wrong not to act in accordance with it. The person answers, "Why should I act

in accordance with it? I have never accepted the principle in the first place." Rawls will say that the reason he should act in accordance with it is because in the original position, he would agree to the principle that people should help each other when it does not involve too much sacrifice for themselves. Naturally, the person's response will be: "But I was never in the original position, and never made such an agreement, so the principle has no bearing on me." Rawls will reply that it does not matter whether or not he was in the original position or had agreed to the principle, as long as the conditions embodied in the original position are reasonable from the point of view of justice, it is his duty to render help to his neighbors when they are in need, since he would agree to the principle of mutual assistance in the original position.

So the way a hypothetical contract "creates" obligations or "generates" moral principles is as follows: (1) The descriptions of an initial situation which are acceptable and reasonable to the contractors who are to choose the principles of justice; (2) The assertion that on the basis of (1) a certain hypothetical contract (which consists of

a set of principles) would be agreed to; (3) The conclusion that the principles reached are the principles of justice, and they have bearing on the contractors. Of course, it needs to be argued or presupposed that the contractors have certain convictions and beliefs as to what conditions embodied in the descriptions of the initial situation are reasonable and acceptable, and it must also be shown that the descriptions do reflect these convictions and beliefs. (We shall come back to this matter later). According to Rawls, (2) can be derived from (1), and (3) from (2). Now for the sake of argument, suppose we agree with Rawls about the logical relations between (1) and (2), and (2) and (3), i.e., from the original position, we can derive a hypothetical agreement, and from the agreement, we are entitled to conclude that the principles described in the agreement are the principles of justice. But if this is the case, we might as well skip (2) and go directly from (1) to (3). The reason we can do this is that the derivation of (3) does not depend on (2), it only depends on (1). (2) is just an intermediate step in the course of the derivation of (3) from (1). It is not an original assumption. Only (1) is needed in the derivation of (3). There are many

different ways to get (3) from (1). To get (3) via (2) is just one of them.

This seems to be true when we re-consider the above example. There, we believe that people should help each other whether or not they have actually agreed to do so. The derivation of this principle of mutual assistance simply depends on the premises we are given in the original position. The statement that we would agree to it does not make the principle one of justice and does not add anything extra to the force of the principle. So it seems the deletion of the notion of a social contract will not affect the derivation of the principle. This is tantamount to saying that, unlike the traditional social contract theory, Rawls' theory is not really contractarian, since, in his theory, the moral principles are not derived from a social contract. Contrary to Rawls' belief, the reference to the notion of a social contract is not indispensable. This result is rather startling. Rawls puts so much emphasis on the notion of a contract; yet, if our analysis is correct, it turns out that the notion does not really play a crucial role in his theory. Naturally, we can expect Rawls to fight back. What responses can we expect

from Rawls? I think he will be insisting on the requirement that the hypothetical contract would be agreed to in the original position by the contractors is a necessary condition for the derivation of the principles of justice. If this is the case, then the reference to the notion of a social contract in his theory becomes indispensable again.

However, this manoeuvre will not rescue him from the criticisms we made. Let us take a look at our arguments again in a more detailed manner. It is a logical truth that if a formula B is derivable from a formula A, and A, in turn, can be derived from a set of formulas C, D, and E, in the course of deriving B from C, D and E, it will be redundant to add A to the original premises, since we can get A from the original premises, namely, C, D, and E. Now, let us use this logical truth to analyze the structure of a hypothetical contract theory. As we saw before, since (3) is derivable from (2), and (2) from (1), therefore, (3) follows from (1). What this tells us is that the derivation of (3) does not depend on (2). If we are given (1), then we will have (3). Of course, we will also have (2). This analysis seems to square with our intuition that the act of agreeing to the acceptance of a certain set of principles as the principles of justice does not make them the principles of justice. What makes them the principles of justice is rather the reasons be-

hind the act of agreeing to this acceptance. The reason we accept (3) is because it follows from (1) but not because it follows from (2). The truth of (3) depends on (1) only.²⁷

Let us consider the problem more concretely in connection with Rawls' theory. We quoted from him before to the effect that many possible interpretations of the initial choice situation are available, and accordingly, there are many different contract theories. Moreover, he says:

We may conjecture that for each traditional conception of justice there exists an interpretation of the initial situation in which its principles are the preferred solution. 28

So if the contractors in the Rawlsian original position choose one set of principles, namely, the two principles of justice Rawls advocates, over another set of principles, say, the principle of average utility, it must be the result of the features embodied in the original position. If we describe the initial situation differently, the contractors might prefer the latter to the former.

The two principles of justice are dependent on or are a function of the original position. The choice itself does not decide which principles are the principles of justice. It is unable to give any principle the status of the principle of justice. It is the reason behind the choice we

made which does the job. The choice is only an intermediate step which does not render any original support to the derivation of the principles in the sense that the principles are derivable from the original position without using the statement that the contractors would make the choice an assumption. If it is not an original assumption, then it is not a necessary condition for the derivation of the principles; hence, it can be deleted. So, again, it seems that the notion of a social contract does not play an indispensable role in Rawls' theory.

There is another way to look at this problem. It has been pointed out by several reviewers of Rawls' book.²⁹ Briefly, since all the contractors in the original position have identical characteristics, there is no way to distinguish an individual A from another individual B there. In the process of deliberation to reach the agreement, there can be no real bargaining among the contractors since nobody knows who he is and what abilities he has, etc. So, it is not a genuine contracting game. It is really the same as the act of rational choice for one person. There are no such things as compromises, concessions, and bargainings involved which are typically connected with the process of making a contract. But in the process of rational choice

by one person, the reference to the notion of a social contract simply does not occur.

Now, we finally turn to the problem of how to interpret Rawls' claim that his theory is to "present a conception of justice which generalizes and carries to a higher level of abstraction the familiar theory of the social contract as found, say, in Locke, Rousseau, and Kant."³⁰

As we indicated earlier, by this claim, we understand as follows: Rawls' theory is to use the contractarian approach to do moral philosophy instead of political philosophy.

The notion of a social contract is employed, in his theory, to derive the principles of justice rather than the principles concerning political authority and obligation.³¹

As we argued before, the traditional theory social contract theory, usually takes the moral principle of fidelity that we should keep our agreement for granted. So, in some sense, moral principles are more fundamental than political principles because the validity of the latter depends upon the validity of the former. When dealing with the problems of political authorities and obligations, the traditional social contract theorists did not question the validity of the moral principles they accepted, whatever these prin-

ciples are. The principles in question are usually derived from some other sources than a social contract. Now, Rawls, in contrast to what most traditional contract theorists did, is trying to derive the higher-level principles, namely, moral principles, via the contractarian approach. This is to say that he is using the contractarian method to deal with moral problems. This interpretation can be corroborated by the following things he says:

In particular, the content of the relevant agreement is not to enter a given society or to adopt a given form of government, but to accept certain moral principles. 32

In order to do this we are not to think of the original contract as one to enter a particular society or to set up a particular form of government. Rather, the guiding idea is that the principles of justice for the basic structure of society are the object of the original agreement. 33

What bearing does this deviation from the traditional use of the notion of a social contract have on the success of Rawls' theory? Is the contractarian approach equally valid for doing moral philosophy as well as social and political philosophy? These are the questions we want to examine next.

Footnotes

1. Rawls, A Theory of Justice, (Cambridge, Massachusetts: Harvard University Press, 1971), Preface, ix.
2. John Rawls, "Outline of a Decision-Procedure for Ethics," The Philosophical Review, 60(1951), 177-79. See also his "The Independence of Moral Theory," The Presidential Address to the American Philosophical Association, Eastern Division, 1974, in Proceedings and Addresses of the American Philosophical Association, 48(1975), 5-22.
3. Rawls, A Theory of Justice, p.11
4. Robert Paul Wolff, In Defense of Anarchism (New York: Harper & Row Publishers, 1970), p.5.
5. What is a social contract is a complex problem. I take it to be a certain kind of agreement we typically make when we are involved in legal matters.
6. This formulation is basically the same as Spencer Carr's which appears in his article "Rawls, Contractarianism, and our Moral Intuitions," The Personalist, 56(1975), 83-95. I made only slight revision. Carr's article is illuminating in distinguishing the structure of Rawlsian contractarianism from that of the traditional social contract theory. However, he fails to recognize one crucial difference between the two; while the traditional theorists used the contract to derive political obligations, Rawls uses it to derive moral principles. This difference is crucial as I shall argue that while the former usage is valid, the latter is not.
7. Plato, Euthyphro, Apology, Crito and Symposium, The Jowett Translation revised and with an Introduction by Moses Hadas, (Chicago: Henry Regnery Company, 1953), p.71.
8. Rawls calls his own theory by this name. To put it briefly, this theory states that in a hypothetical situation of choice, given that we acknowledge certain premises there, the contractors will agree to certain principles as the foundation of their association. (In Rawls' case, they are the principles of justice.) This theory does not commit to the existence of any factual contract, either actual or tacit., I will discuss this kind of contract theory in detail in the following sections.

9. The hypothetical version poses some problem. We will come back to it later when we discuss Rawls' theory.
10. The question what other obligations than the political ones we have is outside the realm of political philosophy. And how are they derived is another matter which we do not have to get into here.
11. There is some difference between a promise and a contract (agreement). For a contract, usually all the contractors will be benefited by entering it. But for a promise, usually only the promisee is benefited not the promisor. But for the moment, we shall ignore this difference.
12. John Searle, "What is a Speech Act?" in Philosophy in America, ed. by Max Black (London: George Allen & Unwin Ltd., 1965), p.224.
13. H.A. Prichard, "The Obligation to Keep a Promise" in his book Moral Obligation and Duty and Interest (London: Oxford University Press, 1968), p.172.
14. Locke argues that for anyone who does not believe in God, the contract dissolves. See David Gauthier, "Why Ought One Obey God: Reflections on Hobbes and Locke," Canadian Journal of Philosophy, Vol. 7, No.3 (1977), pp.425-446.
15. "In the seventeenth century doctrine, rights were attributed to individuals as if they were intrinsic properties of men 'as men.' 'Thus we are born free as we are born rational,' said Locke. Whatever rights are granted a man as citizen of this or that state, his natural rights go with him wherever he goes; they are said to be 'inalienable,' 'imprescriptable,' 'indefeasible.' No government or positive law can deprive him of them, nor can any higher claim prevail against them; if they are to be limited at all, it is only by the consent of their possessor. The contract theory of government was thus inextricably bound up with natural right theory." Natural rights are in turn based on and derived from natural laws. S.I. Benn and R.S. Peters, Social Principles and the Democratic State (London: George Allen & Unwin Ltd. 1959), p.96, See also pp.318-331. See also Gauthier's article cited in the previous footnote.

16. Rawls, op. cit., p.11.
17. Ibid. p.11.
18. Ibid. p.13.
19. Ibid. p.21.
20. Ibid. p.12.
21. Ibid. p.121.
22. Ibid. p.12.
23. Ronald Dworkin, "The Original Position," in Reading Rawls, ed. by Norman Daniels, (New York: Basic Books, Inc., 1975), p.18.
24. Rawls, A Theory of Justice, p.18.
25. Ibid. p.120.
26. Ibid. p.587.
27. Marcus Singer argues in a similar manner when he says:
"We can choose to act on, or abide by, certain principles, and in that sense can choose the principles on which we act, but we cannot, in the original position or out of it, choose that the principles we act on or choose to act on be right or valid. We cannot, in other words, make them right by our act of choosing. For if the principles of morality could be chosen, then whether an action is right or an institution just would be a matter of choice, and it is not. We can acknowledge that some practice is just, but our acknowledgment is not what makes it so." See Marcus Singer, "The Methods of Justice: Reflections on Rawls," The Journal of Value Inquiry, Vol. X, No.4 (W.1976), p.294.
28. John Rawls, op. cit., p.121.
29. For example, Spencer Carr and Marcus Singer. See their articles cited above.
30. Rawls, A Theory of Justice, p.11.

31. In this respect, Rousseau is an intellectual ancestor of Rawls'. According to Rousseau's theory, before people entered into a social contract, they were uncivilized, amoral, etc. It is the contract which elevates people from the uncivilized to the civilized, and transforms them from immoral to moral. See Jean Jacques Rousseau, The Social Contract, Translated with Introduction by G.D.H. Cole, (London: J.M.Dent and Sons Ltd, 1913), Book I, Chapter: III and Book II, Chapter VII.

32. John Rawls, op. cit., p.16.

33. Ibid. p.11..

Chapter 2 Is the Contractarian Method Valid for Doing Moral Philosophy

I. Social Contract, the Institution of Promising and the Obligation to Keep One's Promise

The question whether the contractarian method is valid for doing moral philosophy, i.e., to derive moral principles, seems, as far as we can ascertain from examining his text, never to have entered Rawls' mind. In claiming that his theory is to generalize and carry to a higher level of abstraction the traditional conception of the social contract, there has never been any doubt in Rawls' mind that there might be a qualitative difference between his theory and the traditional theories; and this difference perhaps will make the contractarian approach valid for the traditional theories but invalid for his. Yet, the importance of this question will not be exaggerated if we claim that the success or failure of Rawls' theory depends on the answer to this question. Since, if the contractarian method cannot be validly applied to the derivation of moral principles, Rawls' efforts in doing so will be in vain.

In this chapter, we shall carefully examine this question. We shall argue that there are at least two reasons

why the contractarian method is not valid for doing moral philosophy. The first one is that at least one moral principle, namely, the principle of fidelity, must be presupposed by the contractors in the state of nature. If this is the case, in view of the definition of "contractarianism," it will disqualify the theory in question as contractarian. Secondly, in using the contractarian method to derive moral principles, we will inevitably face a dilemma. If we include among the conditions of the state of nature certain moral principles, then some moral principles are not the result of the contract; therefore, the theory is not fully contractarian. On the other hand, if there is no moral principle presupposed in the original position, how are we to know that the principles agreed to by the contractors are moral principles?

Having made clear in the last chapter the contrast between traditional contract theories and Rawls' theory, it seems natural to ask whether this difference between them has any bearing on the validity of the application of the contractarian method to their respective enterprises. The traditional contract theories were all attempts to derive principles related to political authority and ob-

ligations. We argued at length there that for the contractarian method to be useful in deriving the principles, a moral framework must be presupposed. The traditional contract theories presuppose two things: (1) the availability of the institution of promising; and (2) the validity of the principle of fidelity. The principle of fidelity is a moral principle which has already been accepted as binding by the contracting parties before they gather together to make a contract. But for Rawls, the content of the contract is not to enter a given society or to adopt a set of principles for political ordering; rather, it is to choose the very moral framework itself. And the principle of fidelity itself is one of the principles that must be chosen, hence, justified, by the contract.² In his article "The Method of Justice: Reflections on Rawls," Marcus Singer makes precisely the same point: "For the traditional contract conception has its setting in a framework of moral principles presumed to be already known or accepted. But a contract conception of the principles of morality themselves would at least appear to have no setting in which to operate."³ The curious thing is that Rawls does not seem to recognize that his theory needs such a framework.

In chapter 1, we pointed out Prichard's difficulty with respect to how a promise can create obligations by itself. His answer is that a promise can create obligations only if a previous and more general agreement had been reached beforehand, saying that all agreements must be kept. But, in recognizing the danger of being involved in an infinite regress if we keep on invoking previous agreements to render a promise valid, Prichard admits that what he has pointed out is only a problem to be considered rather than a solution. He says,

But, of course, it would be more accurate to say that what I am suggesting is not a conclusion but a problem for consideration, namely, what is that something implied in the existence of agreements which looks very much like an agreement and yet, strictly, cannot be an agreement.

Rawls, while acknowledging that he follows Prichard's account largely in regard to the issue of how promises can create obligations, thinks he can get out of this predicament via his hypothetical contractarian approach. He claims that through the hypothetical contract theory, the infinite regress can be stopped, and the thing Prichard was looking for which is implied in agreements can be identified. In his own words, "Thus what corresponds to the something,

which to Prichard looked like a prior agreement but is not, is the just practice of giving one's word in conjunction with the hypothetical agreement on the principle of fairness." ⁵

We shall argue that Rawls' manoeuvre will not only be unable to stop him from the infinite regress, but will also undermine his whole enterprise of hypothetical contractarianism. The reason for making this claim is that, while in attempting to derive moral principles, contract theory must presuppose two things: 1) the institution of promising; 2) the validity of the principle of fidelity. In other words, (1) and (2) are more primitive than the notion of a social contract, in the sense that the validity of the latter depends upon the acceptance of the former by the contractors. For Rawls, the contrary is true. He thinks that (1) and (2) are derivable from a contract.

Rawls begins by stating that there is a tendency to conflate the institution of promising and the principle of fidelity. (This principle is a special case of the principle of fairness.). This is a mistake, he claims. The institution of promising is a constitutive convention, which consists of a set of rules. The most basic rule is the one which governs the use of the words "I promise to do X." The rule is roughly as follows: if one utters the

words "I promise to do X" in the appropriate circumstances; one is to do X, unless certain excusing conditions obtain. This is the most fundamental rule of the institution of promising. It is not a moral principle, but a rule of the constitutive convention called promising. In this respect, it is not different from legal rules or rules of games. The principle of fidelity is not a constitutive convention but a moral principle. After this distinction is made clear, Rawls notes two things: First, in agreement with Prichard, that the institution of promising does not by itself create obligations because the contract theory holds that all obligations are derived from moral principles. Thus, he says,

First of all, as the discussion of promises illustrates, the contract doctrine holds that no moral requirements follow from the existence of institutions alone. Even the rule of promising does not give rise to a moral obligation by itself. To account for fiduciary obligations, we must take the principle of fairness as a premise. Thus, along with most other ethical theories, justice as fairness holds that natural duties and obligations arise only in virtue of ethical principles. 6

Secondly, Rawls does not consider promising as a practice which is just by definition. A promise is binding only if it is made under the appropriate circumstances as defined by a set of principles of justice.

So what makes a promise binding is the combination of the principle of fidelity and the fact that the promise has been made. A further provision is that the institution of promising is just according to a set of principles of justice. Since promising by itself does not create obligations, the principle of fidelity will be eventually responsible for the imposition of obligations on the person who has made the promise.

But what makes the moral principles (the principle of fidelity in this case) themselves valid? How are we to justify these principles? Rawls gives a full-blown contractarian answer to these questions: these moral principles are those that would be chosen in the original position. So, we can see that Rawls' position with respect to the problem of how promises can create obligations is somewhat similar to Prichard's. Both of them maintain that a promise by itself alone is not sufficient for creating obligations. The claim that a promise has bearing on the promiser is based on the conviction that a more general agreement that we should keep our agreements exists. For Prichard, this will create some difficulty (because of the infinite regress). For Rawls, this difficulty can be overcome because of the

availability of the principle of fidelity. And this principle itself is, according to Rawls, derived from a hypothetical contract.

At the first glance, this seems unproblematic. The "something" Prichard was looking for so hard yet could not find, can be clearly identified as the principle of fidelity. And we can show the validity of this principle without getting ourselves involved in an infinite regress because it is based on a hypothetical contract. The hypothetical contract itself is supposed to be ultimately responsible for all the moral principles in Rawls' contractarian theory. But upon second thought, this is not so unproblematic. This is so precisely because of the reason we presented earlier in connection with our discussion of the traditional social contract theories, that any form of classical contractarianism must presuppose two things: (1) that the institution of promising is available and that it is a just practice; (2) the validity of the principle of fidelity. The shift from classical contractarianism to hypothetical contractarianism will not make any difference as far as these requirements are concerned, since in making the notion of a social contract the foundation of political or

moral principles, it is a prior logical requirement that the contract itself is binding, for otherwise the purpose of contractarianism will be defeated. But how can we know that a contract is binding if the principle of fidelity itself (which states that we should keep our agreements or promises) is not presupposed? For the classical contract theories, the principle is a moral principle. When they use the notion of a social contract to derive political principles, it does not create difficulties as long as they acknowledge the validity of the principle of fidelity, and try to argue for its validity on the basis of something other than a contract. But for Rawls, since he attempts to derive all moral principles from a social contract, difficulties arise as to the foundation of the principle of fidelity which is a principle whose validity must have been accepted or acknowledged before any contract has been made. Now let us see why Rawls' attempt to use the notion of a social contract to derive moral principles does not work, and what kind of difficulties Rawls will face.

(1) In making a contract, it is typical that the contractors will utter the words that "I agree to the following terms" or "I agree to perform such and such an act"

(or something equivalent to it). In giving our word that we agree to certain terms, we are making a promise that if certain circumstances arise, we shall do what is required by the terms. Usually in making an agreement, we also expect other contracting parties to give their pledge for the terms or for acting in certain ways, though not necessarily for performing the same act as we do. So, an agreement is a form of promise. It is a two-way promise in the sense that all contractors are both promisors and promisees at the same time. It is obvious that the parties to the contract in the original position must have in their hands the institution of promising in order to bring about a contract. If the institution is not available to the contractors, they would not be able to understand what it means to make a contract with other people. It becomes meaningless for them to utter the words "I'll give you my word" or something to this effect. Now, for a promise to have bearing on the promisor, according to Rawls, we must further presuppose that the institution of promising is a just practice. This is where difficulties occur. In the original position, contractors do not have a conception of justice. And according to Rawls, a practice is just

only if it is based on the ground of the principles of justice which are yet to be derived from a contract. But in the original position, the principles of justice themselves are not available yet. This leads us to the question: how are the contractors supposed to know whether the institution of promising is a just practice or not? As a matter of fact, they have no way of knowing it at all. But for Rawls, this requirement that the institution of promising must be just cannot be eased since he does not consider the institution of promising to be just by definition. It is false to say that all promises have the same status. Promises made under certain circumstances will not have bearing on us. For example, we do not want to consider promises forced upon us by threat to be binding; nor do we want to consider valid those promises made when we are under the influence of drugs. Furthermore, if we ease this requirement, the qualifying phrase "the appropriate circumstances" in the basic rule governing the practice of promising becomes unnecessary. Since if we do ease the requirement, it will follow that promises made under drastically different circumstances will not make any difference for the parties involved, i.e., the promisee

and the promisor. A promise made under force will equally specify what a promisor will do as a promise made under the appropriate circumstances. But this seems to contradict our notion of the institution of promising. For a promise to be binding, we think that it should be made voluntarily. This is why Rawls claims that the restrictions on the appropriate circumstances are necessary in order to guarantee the validity of the promises made. But, if this is the case, we have to be able to show that the institution of promising is just. This really creates a problem for Rawls, because the contractors in the original position do not know whether or not the practice of promising is just because they can have this knowledge only if they know the following two things: (1) the institution is just according to the principles of justice; (2) what the principles of justice are. But in Rawls' theory, the principles of justice are the result of the contract; contractors in the original position do not have a conception of justice.

(2) With respect to the principle of fidelity, it is equally apparent that the contractors must have already accepted or acknowledged its validity before an agreement is

reached in order to make the contract itself binding. If the principle is not presupposed, there will be no way of justifying that the agreement reached must be kept, since promise alone does not create obligations. Some contractors who participate in the joint adventure with a bad faith think they can legitimately ask the question when they didn't keep their end of the bargain: "Why should I keep my agreement? There is no moral principle which says that it is wrong for me to break an agreement?" Even if we know they are pretending to be moral, it seems that we have no way of condemning them because the principle of fidelity is not available. People like Hume and Prichard can also raise the kind of philosophical question: "How can promises (contracts) create obligations?" Only if we have presupposed in the original position that the principle of fidelity is valid can we tell them that it is wrong because in the beginning we have accepted the principle that promises (agreements) must be kept. It is clear Rawls himself has made this assumption in the original position:

There is one further assumption to guarantee strict compliance. The parties are presumed to be capable of a sense of justice and this fact is public knowledge among them. This condition is to insure the integrity of the agreement made in the original position ...

It means that the parties can rely on each other to understand and to act in accordance with whatever principles are finally agreed to. Once principles are acknowledged the parties can depend on one another to conform to them. In reaching an agreement, then, they know that their undertaking is not in vain; their capacity for a sense of justice insures that the principles chosen will be respected. 7

What this passage suggests is that we must presuppose, for the purpose of the original agreement, that the contractors have a sense of justice, and this sense of justice will give us guarantee that the agreement reached by the contractors will be honored and respected. Otherwise, the effort by the contractors might be in vain. Rawls' assumption that the contractors in the original position are capable of a sense of justice is only another way of saying that the contractors accept the principle of fidelity in the original position as binding. The sense of justice is just a disguised form of the principle of fidelity itself. To say that the parties can rely on each other to understand and to act in accordance with whatever principles are finally agreed to is the same as saying that contractors must keep their agreements. The principle of fidelity, as Rawls acknowledges, is a moral principle.

In dealing with the issue of how promises can create obligations, the traditional social contract theories invoke

the natural law theory. According to them, political obligations are all derived from a social contract, and the validity of the social contract is based on a personal obligation that we should keep our agreements. This personal obligation is said to stem from the natural law. And the natural laws govern the realm of morality. This is the reason why we said earlier, following Marcus Singer, that the traditional contract theories have their setting in a framework of moral principles presumed to be already accepted. But for Rawls, because he attempts to derive all moral principles from a social contract, this setting is lacking. Yet, as has been argued, he needs one. In order for the contract to be binding, it is necessary that the principle of fidelity must have already been presupposed and accepted in the original position. This means that the validity of the principle cannot be explained or justified by the notion of a social contract. It is rather the other way round. Rawls' attempt to justify and derive the principle of fidelity from a social contract really puts the cart before the horse. According to the definition of "contractarianism" worked out earlier, we can say that a moral theory is contractarian if in that theory all moral

principles are derived from a contract." But, since any form of contractarianism, hypothetical or classical, must presuppose the principle of fidelity, and the principle itself is a moral principle, it follows that no moral theory can be contractarian in the full sense of the word. What this amounts to saying is that the contractarian method is not adequate for doing moral philosophy.

II. A Dilemma

The second difficulty the contractarian approach will face when applied to the derivation of moral principles has its analogue in the traditional social contract theories. The traditional contract theorists begin their theories with the device of the state of nature. What this theoretical device serves is to distinguish between the natural and the contingent. It is said that the conditions embodied in the original contractual situation reveal what is humanly essential rather than express what human beings have become through the course of historical contingencies.^{7a} It is believed that the principles derived from what is humanly essential will carry more weight and have more universal

applicability than the principles adopted because of certain contingent circumstances.

This immediately creates a dilemma for the traditional social contract theories. How do we determine which features are humanly essential so that they should be included in the description of the state of nature and which features are historically or socially contingent to human beings so that they must be excluded from the state of nature? The first horn of the dilemma is that we include too little in the original contractual situation so that the device of the state of nature is not powerful enough to generate a set of desired principles upon which a society can be built. The second horn of the dilemma arises out of including too much among the humanly essential. In doing this, there is the imminent danger that certain conditions peculiar to a given historical period and social circumstances will be mistakenly included in the state of nature as humanly essential. We will run the risk of taking what is historically contingent as universally valid. It, of course, will serve the purpose of enabling us to establish a set of principles upon which a society can be built; but it faces the danger that the principles will not be universally applicable; hence, its validity will be greatly restricted.

This is the kind of difficulty C.B. Macpherson points out in connection with his discussion of Hobbes' theory. He contends that if we interpret Hobbes' psychological propositions about the humanly essential simply to mean those propositions about sense, imagination, memory, reason, appetite, and aversion, then these psychological propositions are insufficient for the derivation of the necessity of a sovereign state. "If on the other hand we use the term psychological propositions to include Hobbes's statement of the necessary behavior of men towards each other in any society(viz., that all men seek ever more power over others),... then the psychological propositions do contain all that is needed for the deduction of the necessity of a sovereign, but they are not about the human animal as such; some assumptions about the behavior of men in civilized society had to be added."

Rawls also talks about the necessity for distinguishing what is natural from what is historically and socially accidental to human beings. Because his contractarianism is an attempt to derive moral principles instead of political and social principles from the state of nature, his emphasis on what is natural is concentrated on what is

humanly essential when human beings are considered as moral beings capable of a sense of justice. The function of the device of the original position corresponds to the state of nature in the traditional contract theories. The conditions embodied in the original position are said to capture what is humanly essential when men are viewed as moral beings. One thing we have to keep in mind is that, when he imposes the various conditions on the original position, the language he uses is loaded with intuitionistic conceptions and suggestions. For example, he often says that a certain condition "looks natural for the contractors to accept," or "the condition seems reasonable to impose," etc. Now, Rawls is very well aware of the dilemma encountered in the traditional contract theories. He understands that if he wants to argue for a set of substantive moral principles, the original contractual situation must contain a set of conditions powerful enough to yield them; but by doing so he takes the risk of importing some controversial assumptions. On the other hand, the conditions imposed on the original position must not be too weak and insignificant to generate the desired principles. So, the dilemma analogous to the one faced by the traditional contract theories

also arises when Rawls sets forth his own contract theory.

Yet the dilemma Rawls faces has its own peculiarity due to the fact that his contractarianism is a moral theory.

What is this dilemma? First of all, let us recall what is a contractarian moral theory. A moral theory is contractarian if, and only if, it asserts that all moral principles (principles of justice about distribution, principles of individual duties and obligations are all included) are derived from and based on contract. Whether or not the contract is actual or hypothetical makes no difference here. According to this definition, if in a theory T, there are certain principles which are not derived from the contract, or, if before the contract is made, the contractors already had some moral convictions and beliefs which cannot be overturned, then T is not a contractarian doctrine. It is because such convictions and beliefs cannot be based on the contract, therefore their basis must be found elsewhere. For Rawls, the first horn of the dilemma is that if a moral theory is contractarian, then, by definition, the conditions describing the state of nature must not include any moral principles or beliefs which have already been accepted by the contractors. In particular,

the contractors are not supposed to be bound to each other¹⁰ by prior moral ties. If this is so, we have no way of knowing why the principles reached by the contractors are moral principles instead of some other kind of principles which have nothing to do with morality whatsoever. In regard to this point, David Lyons has suggested that Rawls' principle of justice can be viewed as a deviation from egalitarianism for the sake of each and every contractor's interests.¹¹ Brian Barry also has his doubts as to whether Rawls has successfully identified the principles of justice by way of a social contract. He says: "But it is important to understand what is claimed to be deducible and what is not. Thus, Rawls does not say that he can give a deductive proof of the proposition that whatever principles would be chosen would be principles of justice."¹²

The second horn of the dilemma is that if we include too much in the conditions of the original contractual situation, in particular, if we include among the conditions that the contractors in the original position have already accepted some moral principles as binding, then the doctrine will be a failure when it claims that moral principles are based on contracts alone. This is so because there are moral

principles which have been accepted before the agreement is reached. In short, it is no longer a contractarian doctrine. This appears to be the case with Rawls's theory. Earlier we have argued that his theory is not contractarian on the basis that the contractors in the Rawlsian original position have already acknowledged certain moral principles as binding, particularly, the principle of fidelity. We even went further by arguing that no moral theory can be contractarian.

Footnotes

1. Rawls has several ways to escape from this delima:
(1) by invoking the notion of pure procedural justice;
(2) by stipulating the role which moral principles play in our society; and (3) by arguing that the conditions imposed on the original position are peculiar to moral theories. We shall deal with these various attempts in the next chapter when we discuss the problem of Rawls' claim that the principles chosen by the contractors in the original position are bound to be principles of justice.
2. Rawls talks about the four-stage sequence of the contracting game, but the primary aim of the contract is to choose the principles of justice themselves. His purpose for talking about the four-stage sequence is to show that the principles of justice chosen in the original positions "define a workable political conception, and are a reasonable approximation to and extension of our considered judgments." See John Rawls, A Theory of Justice. (Cambridge, Massachusetts: Harvard University Press, 1971), pp.195-201. For our contention that the content of a Rawlsian contract is not to enter a given society but to choose the principles of justice themselves, see the two passages we cited in the footnotes 32 and 33 of chapter one.
3. Marcus Singer, "The Method of Justice: Reflections on Rawls,") The Journal of Value Inquiry, Vol. X, No.4 (W 1976), p.289.
4. H.A. Prichard, "The Obligation to Keep a Promise," in his book Moral Obligation and Duty and Interest (London: Oxford University Press, 1968), p.179.
5. John Rawls, op.cit., p.350.
6. Ibid. p.348.
7. Ibid. p.145.
- 7a. For this point, see Milton Fisk's perceptive article "History and Reason in Rawls' Theory" in Reading Rawls, ed. by Norman Daniels, (New York: Basic Books, Inc., 1975), pp. 53-80.
8. C.B. MacPherson, The Political Theory of Possessive Individualism (London: Oxford University Press, 1962), p. 18. According to MacPherson, Locke also committed the fallacy of taking certain features peculiar to the existing society he happened to live in as universal of all human societies. (see p.228 of his book.)

9. We ignore for the moment that it is an impossibility for a moral theory to be contractarian for the reason we offered earlier, i.e., the principle of fidelity must be presupposed for the contractarian approach to moral problems.
10. Rawls states this clearly: "A further assumption is that the parties try to advance their conception of the good as best they can, and in attempting to do this they are not bound by prior ties to each other." See A Theory of Justice, p.128.
11. Lyons suggests something as follows: Suppose that a society has been organized in accordance with the egalitarian principle. Now its members realize that if inequalities to a certain extent are allowed, all of them will be better-off materially. However, they still wish to abide by the egalitarian principle. It seems that from the point of view of self-interest, they are irrational. But there seems to have no reason to consider their society unjust or defective from the moral point of view. Yet, according to Rawls, their society is defective. See his article "Nature and Soundness of the Contrast and Coherence Arguments," in Reading Rawls, ed. by Norman Daniels, (New York: Basic Books, Inc., 1975), p.152.
12. Brian Barry, The Liberal Theory of Justice (London: Oxford University Press, 1973) p.11.

Chapter 3: The Justifactory Use of the Social Contract, Social Contract as the Justification of Moral Principles

I. Social Contract and the Justification of Moral Theories

When a contractarian moralist like Rawls faces the dilemma set out in the previous chapter, how can he proceed? There are two possibilities of escaping from the dilemma: First, he can renounce his allegiance to contractarianism and acknowledge that the description of the state of nature in his theory does include certain conditions which have moral force. By this acknowledgement, it follows that the validity of the moral principles derived from the original contractual situation will not depend on the notion of a social contract alone. This seems to be the case with Rawls' theory notwithstanding his assertion to the contrary. The arguments we have presented earlier that the conditions embedded in the original position carry the ultimate burden for the derivation and justification of his principles of justice bears witness to the truth of this claim. Secondly, he can pledge his loyalty to contractarianism and find some other ways to identify the principles of justice. This is

to say, a contractarian moralist can argue that even if a contractarian moral theory should not import any moral principle into the description of the state of nature, this does not entail that we have no ways of knowing that the principles chosen by the contractors are moral principles since there are other methods to identify the principles of justice. If a contractarian is successful on this count, he will be able to escape from the dilemma by grasping the first horn. This is what Rawls tries to do. In this chapter, we shall discuss the way the dilemma arises in Rawls' theory and his way out of it.

The issue concerns the validity of the contractarian strategy for the justification of moral principles. We are interested in the plausibility of the claim that any principles resulting from a social contract are bound to be moral principles, or principles of justice. With regard to Rawls's theory, we want to ask specifically: is there any reason for us to believe that the principles chosen by the mutually-disinterested contractors concerned to advance their own self-interests, even though they are situated in the highly specified Rawlian original position, are necessarily the principles of justice. Does the

fact that the principles would be thus chosen by the contractors in the Rawlsian original position confer the status of principles of justice on them? This question must not be confused with the question of how we know whether Rawls' principles of justice (suppose we know they are moral principles) are better than other principles of justice. The question we are concerned with is how we know whether the principles chosen by the contractors in the original position have anything to do with morality or justice. The other question concerns the problem of how to rank the various conceptions of justice given that we already know that they all have something to do with morality or justice. The answer to this question depends on the answer to the question we are concerned with, in the sense that only when we are able to determine that all the various theories we intend to compare have something to do with morality. The way to answer the second question, i.e., why theory A is superior* to theory B, might be drastically different from the way for answering the first question. Rawls thinks that his contractarian method is able to provide answers to both of these questions at the same time. It is his contention that the fact that a set of principles would be chosen by the contractors in the original position will

show: 1) the principles are moral principles; 2) the principles are better than other candidates. We shall not be concerned ourselves with the question of whether or not we can rank principles in this way. It suffices to note that the reason a set of principles would be chosen over other candidates is dependent on the specific conditions describing the original contractual situation. Hence, how we rank theories eventually depends on how we rank the various descriptions of the state of nature. We shall concentrate on Rawls's first claim, i.e., any principles chosen by the contractors are principles of justice.

The purpose of Rawls's theory is to provide an alternative conception of justice to utilitarianism, intuitionism, and perfectionism. In view of Rawls' conviction that utilitarianism has dominated the Western ethical tradition for a long time, he must be well aware from the beginning that his theory will give rise to controversies. Therefore, it needs strong justification.

There are, at least, two different sorts of strategy for justifying moral theories commonly employed by moral philosophers. Rawls considers them both as unsatisfactory. The first kind of strategy widely used is what Rawls calls

"Cartesianism." The advocates of this kind of strategy for the justification of ethical theories start with a set of principles which they think are self-evident. And from these self-evident principles, a set of substantive moral principles can be derived by using the technique of deductive reasoning. According to the Cartesians, there is a set of principles which can be seen to be true without the need of proof, and some even think that principles of this sort are necessarily true. The second kind we often encounter in moral philosophy is what Rawls calls naturalism. The naturalist's strategy begins with the attempt to define moral concepts in terms of presumptively non-moral ones, and then proceeds to validate moral principles via the use of the accepted procedure of common sense and the sciences. Although naturalists do not consider first ethical principles as self-evident, they do not consider that their justification poses special difficulties either. As long as the definitions for the moral terms work out, the validation of moral principles can be carried out in the same manner, as it is done in other scientific theories. Rawls rejects both strategies as unsatisfactory. Cartesianism fails because there is no set of substantive first principles which

is self-evident, and which therefore can be used to carry the burden of justification in the manner construed by the Cartesians. As a matter of fact, Rawls thinks that all moral principles are contingent in the sense that they are chosen by the contractors in the original position in the light of general facts. Naturalism is also doomed because it makes definitions the major part of an ethical theory, and because of this, they in turn need to be justified. The reason Rawls rejects these two strategies is not hard to find if we recall what Rawls' conception of moral theories is and how deeply he has been influenced by Quine. He contends that in order to construct a substantive moral theory, the method of linguistic analysis is inadequate. In his article "On Philosophic method," Kai Nielsen puts this point succinctly:

In thinking about the work of Rawls in particular, it is important to keep in mind that he has been deeply influenced on foundational matters by his colleague Quine. In particular, this means that he does not attach philosophic significance to the distinction (putative distinction) between the analytic and synthetic and he does not regard it as his proper philosophical task to give an analysis or an explication of moral and political concepts. Making no sharp distinction -- indeed regarding such a distinction as artificial -- between, on the one hand, the analytical task of the explication of concepts and, on the other, an examination of substantive

matters, let alone regarding it as the only proper philosophical task to do the former; 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. 1

In view of this, Rawls must have his own method for the justification of moral principles. Basically he employs two different kinds of method for both the derivation and justification of moral principles. According to Rawls, a set of moral principles can be justified in the following two ways: (1) it can be shown that the set of principles would be chosen by the contracting parties in the original position; (2) it can be shown that the principles agreed upon match our considered convictions of justice or extend them in an accepted way.² We shall put off the discussion of the method of reflective equilibrium for the justification of moral principles later.

There are two basic ideas shared by all contractarians. The first is the notion of collective choice. According to the contractarians, the reason a set of principles is justified is because the principles have been or would be chosen by the contracting parties. This is certainly what

Rawls has in mind when he uses the contractarian strategy for the justification of moral principles. We have already offered some arguments why this strategy is not valid for justifying moral principles. The second basic idea all contractarians hold is the deep-rooted suspicion of intuition. They want to reduce our reliance on intuition to the minimum in matters of the choice of political or moral principles even though they admit that we cannot dispense with our intuition altogether. The way they set up their theories is, first, to stipulate a set of conditions describing the original contractual situation, and secondly, to derive a set of principles from it by strict deductive method. This is exactly the procedure Rawls follows in setting up his theory.

The argument aims eventually to be strictly deductive...we should strive for a kind of moral geometry with all the rigor this name connotes. 3

Even with this goal in his mind, Rawls still has to admit that he relies heavily on intuition in the course of constructing his theory. The sentence immediately follows the above quotation reads, "Unhappily the reasoning I shall give will fall far short of this, since it is highly intuitive throughout. Yet it is essential to have in mind the ideal one would like to achieve."⁴

II. Rational Choice and Moral Principles

How is Rawls to achieve the deductive rigor he strives for? The strategy Rawls uses is to provide a way of transforming the basic moral problems, for which there is no generally recognized procedure to follow, to a more manageable problem of rational prudence, and thus to replace controversial moral judgments by judgments of rational prudence.⁵ Thus, the fundamental problem of adopting and justifying a set of moral principles becomes the problem of how to find a solution to a situation of rational collective choice.

Rawls sets up this choice situation in the following way. We are asked to imagine a hypothetical situation which is described as the original position in which a group of people are to launch a joint adventure. They realize that subsequently they can all benefit from this cooperation. The purpose of this joint adventure is to search for a set of principles which will assign duties and rights to each and every contractor in the actual society. There is a set of conditions imposed on the original position which Rawls thinks would be fair to each and every contractor as free and equal participant in the adventure. The con-

ditions are divided into two parts. On the objective side, we are told that the condition of relative scarcity of resources obtains, and cooperation is both necessary and possible. On the subjective side, the contractors are conceived to be rational in the standard economic sense of taking the most effective means to ends, of being self-interested, and of being temporarily behind a veil of ignorance. The veil of ignorance prevents them from knowing who they are, what their talents, tastes and religious beliefs are, and so on ; yet it does not deprive them of any access to theoretical knowledge. Rawls puts special emphasis on the features of the relative scarcity of resources and the self-interestedness of the contractors. This latter feature entails that the contractors in the original position have no moral motivation such as benevolence, good will, respect for persons, etc. The only thing they are concerned with is the advancement of their own interests (even though they do not know who they are in the actual society). In other words, they are rational egoists despite Rawls' belief to the contrary. Rawls contends that it is a mistake to consider the contractors in the original position as egoists. His contention is based on the belief that whether or not a person is egoistic depends on what

sort of ends he pursues. If wealth, position, power, social prestige are a person's final ends, then he is an egoist; if his aim is something other than this sort, he is not an egoist. It seems to me this is contrary to the way we usually understand what an egoist is. A person is egoistic if what he is interested in is solely the realization of his own ends, no matter what these ends are (apparently it has to exclude the possibility that he is concerned with other people's welfare), in other words, he is totally lack of moral motivation. Whatever he does, he does it only for his own good. An artist whose only goal is his own achievement in art is no less egoistic than a politician whose only end is to gain power or a businessman whose only concern is the accumulation of wealth for himself. It seems that the contractors in the original position are egoists in this ordinary sense of the word as we commonly understand it.

Now, what has been presented is in reality clearly a problem of rational prudential choice. The contractors so situated, who are totally unmotivated by moral considerations and whose only concern is the maximization of their own interests, are trying to reach a solution for

defining the basic terms of their association. The solution consists of a set of principles. Rawls makes the following claims: (1) there is a solution to their problem of rational choice; (2) any principles agreed upon by the contractors are moral principles, or, in Rawls' case, principles of justice; (3) the fact that a set of principles would be chosen by the contractors constitutes the required justification for the claim that these principles are moral principles or principles of justice; (4) the fact that a set of principles would be chosen over other candidates proves conclusively that one conception of justice is superior to another.

As to (1) and (4), Rawls thinks that: (a) his two principles constitute the solution to the choice problem; and (b) the fact that his two principles would be chosen proves that his conception of justice is superior to other alternatives, for example, perfectionism. Whether these two claims are valid does not concern us here. What we are interested in is (2) and (3). These two claims are closely related. Rawls' contractarian approach raises the following doubts: Is there any reason for us to believe that the solution to a rational prudential choice problem

constitutes a set of moral principles? What reason is there for us to believe that any principles which would be chosen by rational egoists as the best way for each to maximize his interests are necessarily principles of justice? Is it legitimate to reduce the problem of the adoption of moral principles to the problem of rational prudential choice? Isn't there a clear distinction between morality and prudence?

Rawls himself is well aware of this distinction, and he also realizes that justice and efficiency are two distinct categories. Some social arrangement may be efficient, but it is not just, or vice versa. He rejects the principle of utility on the ground that it sometimes requires lesser life prospect for some members of society in order that the average or total utility can be increased. This may be efficient from the economic point of view but it is unjust from the moral point of view. According to Rawls, for utilitarianism, "the correct decision is essentially a question of efficient administration."⁶ If this is the case, the principle of utility is a principle of efficiency with respect to the social arrangement in the course of the production of utility. It has nothing to do with morality

or justice just like traffic rules have nothing to do with justice. This statement can be further corroborated by the contrast Rawls makes between his theory and utilitarianism. He says, "Thus while the contract doctrine accepts our convictions about the priority of justice as, on the whole, sound, utilitarianism seeks to account for them as a socially useful illusion."⁷ The statement implies that moral principles only play a subordinate role in utilitarianism, and when they are in conflict with the principle of utility, it is always the latter which merits our considerations. Undoubtedly, Rawls believes that his two principles would be chosen over the principle of utility by the contractors in the original position. But the important thing is that he does not rule out the possibility that the principle of utility might be preferred to his own principle. He says:

To be sure, it has not been shown by anything said so far that the parties in the original position would not choose the principle of utility to define the terms of social cooperation ...It is perfectly possible, from all that one knows at this point, that some form of the principle of utility could be adopted, and therefore that contract theory leads eventually to a deeper and more roundabout justification of utilitarianism.⁸

Rawls' claim that the principle of utility is essentially a principle which concerns the question of efficient

administration implies that the solution to the problem of the rational prudential choice as defined by Rawls is not necessarily a set of moral principles. It might be a set of principles which is only concerned with the efficient arrangement of social institution. This raises serious doubt about Rawls's claim that any principles chosen by the contractors in the original position are moral principles or principles of justice.

Now let us see why Rawls thinks the social contract can play a justificatory role in moral philosophy. His contention that any principle chosen by the contractors is principles of justice is derived from his conviction that the principles of justice can be regarded as emerging from a fair procedure. The function of the original position is just to guarantee that the procedure of the choice is fair. "The idea of the original position," Rawls maintains, "is to set up a fair procedure so that any principles agreed to will be just. The aim is to use the notion of pure procedural justice as a basis of theory." ⁹ The justification behind the use of pure procedural justice lies in the conviction that in dealing with the problem of distributive shares, besides the procedure itself, there is no independent criterion of just distribution that we can rely on.

The notion of pure procedural justice is best understood by a comparison with perfect procedural justice. Rawls has provided us with an example of perfect procedural justice. A number of people are to divide a cake. Assuming that the just division is of an equal share for everyone. How are they going to achieve this? The obvious answer is to have a person divide the cake and get the last piece after others have made their pick. Of course, we have to make a number of assumptions, for example, that there are no technical difficulties in dividing the cake equally, and that the person who divides it wants for himself the largest share possible, etc. Two features stand out in this example: (1) there is an independent criterion for what is the just division, that is, the criterion is defined separately from and prior to the procedure which is to be followed for achieving it; and (2) a procedure is available for the cake-eaters to bring out the desired outcome. Since in this case, there is an independent criterion for what is the just division of the cake prior to the procedure, the procedure of rational prudence designed for achieving this outcome clearly does not justify the criterion even though the criterion and the outcome of following the procedure coincide. As a matter of fact, it is rather the other way round. Further,

if different criterion has been assumed in the beginning, for example, to each according to his needs, the outcome brought about by the above procedure will not coincide with it, and hence, it becomes unjust. In contrast to the perfect procedural justice, "pure procedural justice obtains when there is no independent criterion for the right result: instead there is a correct or fair procedure such that the outcome is likewise correct or fair, whatever it is, provided that the procedure has been properly followed."¹⁰ Rawls uses gambling to illustrate this idea,

If a number of persons engage in a series of fair bets, the distribution of each after the last bet is fair...Whatever this distribution is. I assume here that fair bets are those having a zero expectation of gain... The betting procedure is fair and are freely entered into under conditions that are fair. Thus the background circumstances define a fair procedure... In this sense, all of these particular distributions are equally fair. 11

Now the basic idea of pure procedural justice is this: since there is no independent criterion of what is a just distribution, it is impossible for us to decide from the beginning what distribution we should strive for because it is just. And because of this lack of objective criterion, what we can do is to set up a fair procedure to guide our deliberation. Since the procedure itself is fair, therefore,

whatever outcome follows from it will be guaranteed to be just provided that the procedure is carried out properly.

The validity of using the notion of pure procedural justice on the problem of the adoption of moral principles is not self-evident. Utilitarianism does not interpret the problem of the choice of moral principles as a scheme of pure procedural justice. It seems at least Rawls has not shown that the utilitarians are wrong in this respect. Aside from this, there are some more difficulties. In order to be able to guarantee that the outcome of the deliberation is just, it is essential that the procedure and the background circumstances must be just. Yet, before a set of principles is chosen by the contractors, we lack a standard of what is just. The principles of justice are supposed to be the result of the deliberation rather than the premises used in the deliberation. In view of this difficulty, Rawls has to retreat a little and settle upon the idea of "fairness." Instead of claiming that the procedure of deliberation and the background circumstances must be just, the most Rawls can ask for is that they be fair. The various conditions imposed on the original position are designed to guarantee just that. For instance, the assumptions that contractors

do not know who they are, what their talents and religious beliefs are, are all designed to guarantee that each and every participant will be equally treated in the original position. In this way, Rawls maintains, the fairness of the background circumstances can be translated into the justice of the outcome. Hence, Rawls claims that any principles reached by the contractors, are bound to be the principles of justice because they are the result of a fair bargain.

The line of arguments does not satisfactorily answer the question we asked at the outset: Is there any reason for us to believe that the solution to a problem of rational prudential choice is bound to be a set of moral principles? We raised the question because we believed that any attempt to use the contractarian method to do moral philosophy will face a dilemma: if a theory is contractarian, then all its moral principles must be derived from a contract, hence, the description of the state of nature must not include any moral principles. But if this is so, how do we know that the principles agreed to are moral principles? On the other hand, if we include certain moral principles in the description of the state of nature, the theory is no longer contractarian, Rawls tries to escape from the dilemma by

grasping the first horn. His strategy is to use the notion of pure procedural justice to define moral principles. If we have good reasons to believe that the principles chosen by the contractors will be principles of justice, as the result of following a fair procedure of deliberation, this still successfully show the validity of the contractarian approach. In so doing, we do not have to import moral principles into the description of the state of nature. Unfortunately, the dilemma is not so easy to escape. In order to use the notion of pure procedural justice, Rawls must employ the notion of fairness to define the procedure and the background circumstances of the choice. Rawls argues that the outcome is reached by following a just procedure if the background circumstances which define the choice situation and the procedure of deliberation are fair. But the notion of fairness itself does not seem to be a morally neutral one. In order to claim that a certain set of conditions which describes the original contractual situation is fair, it seems we have to presuppose certain principles to support our claim, and these principles are not morally neutral. This can be shown by what Rawls says about the purpose of the device of the original position:

The aim is to rule out those principles that it would be rational to propose for acceptance, however little the chance of success, only if we know certain things that are irrelevant from the standpoint of justice. 12

This passage shows how the contractarian approach to moral philosophy is committed to a circularity. One of the most basic problems for moral philosophy is to characterize the moral point of view. In designing the device of the original position, we are supposed not to have an idea of what is just or unjust. According to the contractarians, the idea is supposed to be the result of the contract. If this is the case, how are we to judge which conditions are relevant and which are irrelevant from the standpoint of justice? This passage implies that to characterize a contractual situation as fair, at least we must know which conditions are irrelevant from the moral point of view, for otherwise it would be illegitimate to rule out the possibility for some principles to be chosen, even if the chance is very little. Therefore, in order to characterize certain conditions as fair, we seem to have to presuppose certain principles of justice beforehand. The Rawlsian original position bears witness to this: why do we consider that no one should be advantaged by natural fortune or social

circumstances as fair? Is this because we have already accepted the moral principle that people are entitled to equality? Ronald Dworkin suggests that there is a deep theory behind Rawls' contractarianism. At the heart of Rawls' "deep theory," according to Dworkin, is the claim that each individual has the right to equal concern and respect. This right, Dworkin argues, is not a product of the contract but a presupposition of Rawls' use of the contract.¹³ It seems that Rawls' attempt to escape from the dilemma by grasping the first horn fails because in trying to do so, he is caught by the second horn of the dilemma.

There is another difficulty with regard to the use of the notion of pure procedural justice in dealing with the choice of moral principles. Even if we accept Rawls' contention that the original position embodies fairness, we can still ask whether the fairness of the background circumstances and the procedure of deliberation are sufficient to guarantee that the outcome is just. The name of Rawls' theory "Justice as Fairness" does not mean that "Justice" and "Fairness" are identical; what it suggests is only that the conception of justice is derived from the fairness of the procedure of choice and the background circumstances.

In a key passage, Rawls explains the propriety of the name "Justice as fairness" for his theory:

The original position is, one might say, the appropriate initial status quo, and thus the fundamental agreements reached in it are fair. This explains the propriety of the name "Justice and Fairness"; it conveys the idea that the principles of justice are agreed to in an initial situation that is fair. 14

The arguments in the passage can be put in the following form:

- (1) C is a set of choice conditions which is fair;
- (2) P is a set of principles chosen by the contractors situated in C; therefore,
- (3) P is itself fair.

It follows from (3) that

- (4) P is set of principles of justice.

There are two questions here which are puzzling. First, the term "Fair" is ambiguous in the passage. Rawls has never clarified it throughout the book. It is ^{ascribed} predicated ✓ to the result of the contract in (3), but in (1) it is predicated to the initial conditions and the process of choice. Rawls has not offered us any argument to show that by using pure procedural justice, the fairness of the procedure of choice and background circumstances can be translated into the result of the deliberation of the contractors. He only

asserts that it will be the case. But this is not as troublesome as the next question. Suppose we grant that fairness can be preserved by valid deductive reasoning so that the fairness of the initial choice conditions will be carried over to the results of the choice. Even if this is true, we are still troubled by the claim that (4) follows from (3). Rawls has not offered any reason for the claim that valid deductive reasoning can transform the fairness of principles to the justness of principles. He admits that by calling his theory "Justice as Fairness," he does not intend to be understood as claiming that these two concepts are the same. But if this is so, what reason are there for us to believe that he has proved that from the fairness of the initial situation and the procedure of choice, the result is bound to be a set of moral principles? It seems that by invoking the notion of pure procedural justice, the most Rawls can show is that the principles agreed to by the contractors are fair, but it takes a lot more arguments to prove that they are principles of justice.

III. The Role of Justice

Rawls' second way out the dilemma is by attempting to specify the role principles of justice play in society.

He believes, if we can successfully identify this role independently of the notion of a social contract, we will be able to escape from the dilemma.

Rawls says in several places that the way we come to know a certain set of principles as principles of justice (or moral principles) is by recognizing the role they play in society. Principles of justice are defined as principles which assign our duties and rights in the event of social cooperation. He says:

A set of principles is required for choosing among the various social arrangements which determine this division of advantages and for underwriting an agreement on the proper distributive shares. These principles are the principles of justice; they provide a way of assigning rights and duties in the basic institutions of society and they define the appropriate distribution of the benefits and burdens of social cooperation. 15

The concept of justice I take to be defined, then, by the role its principles in assigning rights and duties and in defining the appropriate division of social advantages. A conception of justice is an interpretation of this role. 16

It follows from this definition that the principles of utility, perfection, etc., are all principles of justice since they all play the role of assigning rights and duties in the course of social cooperation.

It seems if principles of justice can be successfully identified or defined in terms of the role they play, Rawls will again be able to escape from the dilemma by grasping the first horn. Identifying the principles of justice in this way, he has no need to invoke the notion of a social contract. This notion is used only in the derivation of the principles in his theory. The question of how we know that the result of the derivation is a set of moral principles will be taken care of by other means, that is, the role they play in a social cooperation. This enables him to say that in his theory, even if no moral principles are presupposed in the original position, he will still be able to assert that the principles chosen by the contractors are principles of justice because he knows the role they will subsequently play in society, that is, to assign our duties and rights.

But this manoeuvre will not help Rawls to escape from the dilemma because the definition for the principles of justice itself is defective. By saying that principles of justice are the principles which assign our duties and rights in social cooperation, he has only provided a necessary condition for a set of principles to be principles of justice; yet, a sufficient condition is still lacking. The reason

for this statement is that there are certain principles which play the same role of assigning duties and rights in a social cooperation, but we are reluctant to identify them as principles of justice. Let us take an example to illustrate this.¹⁷ Suppose for usual prudential reasons of economy, two friends and I decide to form a car pool. We agree that Monday will be my turn to drive. By agreeing to this arrangement, I have agreed to the rule that I should drive my friends to work on Mondays; this agreement makes the fact that today is Monday a reason for me to drive. We can further suppose that my choice of Monday is not arbitrary; it has good prudential reasons. The rule that I should drive on Mondays and other people on other days is based on a contract we have all agreed to and on the rules which assign duties and rights to each and every contracting party. However, even if the rules assign duties and rights to us, the car poolers, it seems quite odd to call them moral rules. The fact that today is Monday does create a reason for me to drive my friends to work, because I have agreed to the rule that I shall drive on Mondays, but the rule itself does not become a moral rule simply because it assigns duties and rights to my friends and me; moreover, the reason that I should drive my friends to work because it is Monday can

hardly be considered a moral reason. There seems to be no reason that we cannot change these rules at any time when we think it is convenient to do so, and we can always trade with each other if we agree. Yet, considering them as moral rules simply because they assign duties and rights, it seems to be contrary to our intuitive understanding of what moral rules are. We can hardly imagine that moral rules are the sort of rules that can be changed so easily simply for the sake of convenience. Of course, in agreeing to the rule that I should drive on Mondays, I cannot refuse to fulfill my obligation for no reason. If I refuse to drive my friends to work on Mondays without any explanation, they can accuse me of violating a moral rule, that is, the principle of fidelity. This principle is a moral principle that we are supposed to acknowledge in the beginning if we intend to enter into an agreement or make a promise. But by my refusal to drive on Mondays, it seems odd for my friends to accuse me of violating a moral rule that I should drive on Mondays even if I have agreed to the rule that I should drive on Mondays.

This discussion shows that Rawls' definition for principles of justice is defective. The definition only stipulates the necessary condition for a set of principles to be

principles of justice, that is, it says that a set of principles can be regarded as principles of justice only if they assign duties and rights in a social cooperation, but his definition does not provide the sufficient conditions. It does not succeed in showing us that a set of principles are principles of justice if they assign duties and rights. Suppose that we know a certain set of principles are principles of justice (actually, of course, we do not have any way of knowing this yet), we are entitled to say that they are concerned with the distribution of rights and duties in the event of social cooperation; but it is false to say that if a set of principles does the job of assigning duties and rights then they must be principles of justice.

IV. The Original Position and the Moral Point of View

Rawls is aware that the optimal solution to the problem of rational prudential choice does not necessarily consist of a set of moral principles. In orthodox price theory, for example, it is claimed that the institution of competitive market, in which all individuals try to advance their own interests and are not concerned with other people's interests, will lead to an equilibrium. This state of af-

fairness is supposed to be optimal to everyone. But nothing assures us that the equilibrium is a just one. It is with this difficulty in mind that Rawls has to find some other way than the contract method to demonstrate that the principles agreed to by the contractors situated in the original position are principles of justice.

Rawls' third way out of the dilemma is to claim that the original position embodies conditions which are peculiar to moral theory. He claims:

The moral assessment of equilibrium situations depends upon the background circumstances which determine them. It is at this point that the conception of the original position embodies features peculiar to moral theory. 18

The claim is tantamount to saying that, in contrast to the theory of price, or other kinds of social theory, the choice reached by the contractors in the original position will be a just one because the conditions embedded in the original position are peculiar to moral theory. This strategy for justifying the claim that principles agreed to by the contractors in the original position are moral principles, is quite different from the contractarian strategy of justification of the claim. The contractarian strategy bases the claim on the contention that any principles

chosen by the contractors as the optimal solution to certain problems of rational prudential choice are principles of justice. Now the new strategy has shifted the burden of justification from the notion of a social contract (an agreement) to the peculiarity of the conditions embodied in the original position. This strategy, as we can see, is radically different from the previous strategy of substituting prudential choices for moral choices. Its success or failure depends on whether the conditions imposed on the original position are really peculiar to moral theory.

In assessing this new strategy, there are two approaches available to us: (1) We can examine each and every condition of the original position and see whether it is peculiar to moral theory;¹⁹ (2) we can also try to argue that one or more conditions in the original position are not compatible with the moral point of view, that is, we can show that the presence of one or more particular condition(s) entails the impossibility for the contractors to contract themselves²⁰ into a morality.

It is the second approach that we will take. The gist of our arguments is the belief that only moral beings with moral values and motivations will accept or agree to moral principles. A group of self-interested people in making

a contract might be able to agree to a set of principles, and these principles might be exactly the same as the ones we consider as moral principles (supposing we have ways to validate them,) but the fact that they are exactly the same is only a coincidence. The fact that a group of self-interested people agree to a set of principles does not constitute any justification for the claim that they are moral principles.

Rawlsian contractors are, by stipulation, mutually disinterested. By this it is meant that they are never motivated by such moral or altruistic considerations as benevolence, good will, respect for persons, etc. We have argued before that they are rational egoists in the sense that the only good they have in their minds is the advancement of their own self-interests. In the original position, their interests are expressed through the desire of obtaining as many primary goods as possible. We can call them perfect egoists in contrast to the egoists we meet in 'ordinary' life; while the latter ordinarily take reasons of self-interest to outweigh moral reasons, the former never act from other motivations than self-interest.

One of the most fundamental tasks of moral philosophy is to characterize the moral point of view. Admittedly, in the beginning, we do not have a clear notion what this point

of view is; but we do have some vague ideas about it. Among these is that it is different from and sometimes in conflict with the point of view of self-interest, that is to say, sometimes in a certain situation, we are required to do certain things, if we take up the moral point of view, but if we take up the viewpoint of self-interest, we will be required to act differently. An extremity occurs when the requirements of the two points of view are incompatible. It sometimes happens that in a certain situation, an immoral act considered from the moral point of view is the most prudential act when considered from the point of view of self-interest. Rawls is very aware of the difference and frequent incompatibility between these two points of view. He admits that egoism taken as a system for action-guiding is logically consistent, but it is incompatible with what we intuitively regard as the moral point of view. He thinks that, "The significance of egoism philosophically is not as an alternative conception of right but as a challenge to any such conception."²¹ This statement implies that anyone who takes up the egoist's point of view is not in the realm of morality.

A person takes up the moral point of view or accepts certain principles as binding when he takes certain facts identified as moral reasons by accepted principles as suf-

ficient for his acting in a certain way. For example, if we accept the principle that all human beings should have equal liberties, the fact that a portion of the human race has fewer liberties than the rest should be a sufficient reason for us to attempt to correct this situation. Similarly, one takes up the viewpoint of self-interest when one takes the fact that certain actions will enhance one's own interests as sufficient reason for one to act in a certain way. An egoist who always acts from his own interests or gives what he takes to be his own interests priority will sometimes act in accordance with moral principles because he thinks such way of acting will serve his own interests; for example, he will get some rewards, or, his reputation will be greatly improved. However, in so acting, an egoist does not act from moral principles. He sees only instrumental values in them. Whenever there is a conflict between what is required by moral principles and the principle of self-interest, he will follow what the latter requires. Of course, for his long-term interests, a calculating egoist might occasionally follow what moral principles require even if it involves certain sacrifices on his part. But he can never see the good of following moral principles as an end in itself. In contrast to this, a moralist acts not merely in accordance with moral

principles, he acts from them. He sees moral reason itself as sufficient for him to act in a certain way. He thinks that to act morally is good in itself, or at least is something that he categorically must do, which does not need any further justification. We can consider an overriding reason as one which is itself sufficient to warrant certain actions. For example, to an egoist, to promote his own interests is an overriding reason on all occasions. This, of course, does not rule out the possibility that occasionally, for his long-term interest, he will act in accordance with what moral principles demand even if he has to sacrifice temporarily. A person who takes up the moral point of view is one who always sees moral reasons as overriding even though to act morally will sometimes require him to make some self-sacrifice. This does not mean that sometimes weightier moral reason will not override a lighter one.

Rawlsian contractors are perfect egoists in the sense that they always take reasons of self-interest as overriding. They act solely from the principle of self-interest under any circumstances. Will such people take up the moral point of view which will sometimes require them to sacrifice their own interests? Let us imagine a group of such people gathering together in deliberating on a set of principles to define

their association. In the first place, the reason they will do this is because from the viewpoint of self-interest, they think that having some principles to govern their association is better than having everybody simply pursue his own interest. This is so because in the original position they know that they are "roughly similar in physical and mental powers; or at any rate, their capacities are comparable in that no one among them can dominate the rest. They are vulnerable to attack and all are subject to having their plans blocked by the united force of others."²² With this understanding, they think that cooperation among them is better for the purpose of advancing their own interests. But this does not mean that in deciding to join the cooperative venture, they are prepared to make sacrifices of their own interests. Their reasoning for participating in the cooperation is as follows: even though the sole purpose of their life is the advancement of their own interests, it is more often than not that this goal cannot be achieved if each and every person is to pursue his or her own interests without cooperation; this is so because very often people's interests will conflict, and when this occurs, if there is no rule or principle they can resort to for adjudicating, the result is damaging to both parties involved. So they figure that it is better

for them, in the pursuit of their interests, to accept some principles which will do the adjudication when conflicts occur. They understand that even though the adoption of these principles will sometimes require them to do things which are contrary to their interest, yet in the long run, everybody will benefit from the acceptance of the principles. An egoist figures that if everyone is to pursue his own interest he might obtain a certain amount of utility for himself. But if everyone follows a set of principles of common choice, the utility he will obtain for himself will be bigger than the one he is able to secure when there is no cooperation, even though the acceptance of the principles occasionally requires him to make some sacrifices. This is how the Rawlsian contractors in the original position figure when they decide to join the cooperative adventure. Obviously it does not entail that when they decide on certain principles to define their association, they suddenly undergo a conversion from the viewpoint of self-interest to the moral point of view. For them, the overriding reason for acting in any situation is still the furthering of their own interests. They might accept certain principles for governing their behavior, but their acceptance of the principle is sole-

ly for the reason of advancing their own interests in the long run. If this is so, we can hardly say that they have adopted the moral point of view since to say that a person adopts the moral point of view amounts to saying that moral reasons are overriding for him, and he needs no further justification besides the moral reasons for him to act in a certain way. They cannot see moral reasons, which sometimes require sacrifices on their part for the good of other people as an overriding reason for acting. Some critics argue that Rawls' theory is egoistic on the basis that the contractors in the original position are egoists. Rawls is right in refuting this argument because to characterize the contractors as mutually disinterested in the original position does not entail that persons in ordinary life are similarly disinterested in one another's welfare. But our arguments that his theory is egoistic does not rest on this confusion. What we have argued is that Rawlsian contractors, characterized as rational egoists, will not accept moral reasons as overriding. The fact that sometimes an egoist accepts certain moral principles and acts in accordance with them does not mean that he is a morally good person. He never considers morality as an end in itself. To follow Kant, we can say an egoist will and can only act in accordance with morality

but never from morality. This bears out the belief we had in the beginning, that is, only beings with moral values and motivations will be able to take up the moral point of view.

Rawls has a clear answer to our contention. He admits that egoists will not accept moral principles as overriding; but he argues that the device of the original position has contained other conditions, and when they are combined, even if the contractors are mutually disinterested, they will be forced to consider the choice of principles from a general point of view. It is the combination of these conditions which prevents the contractors from choosing the principles to accord with their own particular interests. He says,

...it is tempting to think that it will not yield the principles we want unless the parties are to some degree at least moved by benevolence,... Now the combination...of conditions forces each person in the original position to take the good of others into account. 23

There are two points we want to make in regard to this reply. First, we have argued before that Rawls' theory is not contractarianism on the ground that the contractors situated in the original position are not really engaged in a contracting game; they are only making a rational choice for one person. This is so because all the contractors have

identical characteristics, and there is no way by which we can distinguish an individual A from another individual B. In the process of deliberation, there is no room for such things as compromise, concession, and bargain, which are typically connected with the process of making a contract. Now, some have argued that the aim of the veil of ignorance is to obtain impartiality. But if impartiality simply means the elimination of egoism, the formal constraints on the concept of right are sufficient to carry out the mission as Rawls himself so acknowledges.²⁴ But if the formal constraints are sufficient for ruling out egoism, what is left for the veil of ignorance to do? I take it that the purpose of the veil of ignorance is to enable the contractors to adopt the maximin rule as the guidance in the process of their deliberation. The veil of ignorance has prevented the contractors from knowing all the things except knowledge of the theoretical kind. It deprives them of the knowledge of who they are, what their conceptions of the good are, and what their talents, tastes, and religious beliefs are. It is to create a situation in which, according to Rawls, the best strategy for making choices is to adopt the maximin rule. The rule tells us that "we are to adopt the alternative²⁵ the worst outcome of which is superior to the worst outcomes of the others."

The maximin rule becomes the best strategy for decision-making when three chief features appear in a situation of choice:

"First, since the rule takes no account of the likelihoods 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." 26

Rawls believes that the original position embodies all these three features. Whether this is really the case does not concern us. Some of Rawls's critics, e.g., Brian Barry, thinks it does not. What we are interested in knowing is what follows from the adoption of the maximin rule. Since the veil of ignorance has deprived the contractors of the knowledge of who they are, they have no way of knowing which principles to adopt in their own interests. It will be futile for us to tell them to make the choice from a general viewpoint because an egoist, by definition, will not take up a universal or general point of view. So, Rawls must provide a way for them to be able to identify themselves temporarily with a certain social group, otherwise, the choice cannot be made. The maximin rule is precisely to achieve this. By not

allowing the contractors to have the knowledge of who they are, Rawls argues, they will naturally adopt a conservative approach to decision-making, and therefore, identify themselves with the least-advantaged group of the social strata. They figure that since it is not likely for them to know the possible circumstances and the probable outcome of their choices, they had better be careful and play it safe. In so deliberating, they think it is in their interests to maximize the interests of the least-advantaged group of the social strata. The adoption of the maximin rule enables the contractors to find temporarily their own identities in the process of making a choice. Again, whether this reasoning is sound does not concern us here. It suffices to note that some critics, for example, Barry, contend that whether the contractors will adopt the maximin rule, and subsequently, identify themselves with the least-advantaged class, depends on their attitude towards risk-taking. But we wonder why the adoption of the point of view of the least-advantaged group represents impartiality. Rawls does not seem to have an answer to this.

Secondly, we know that one of the aims of the device of the original position is to obtain impartiality. But it seems that impartiality is only a necessary condition for the adoption of the moral point of view instead of a sufficient

one. If we know that someone takes up the moral point of view, then we can expect him to be impartial, but we cannot infer from this that if he is impartial, then he has adopted the moral point of view. Let us use an example to illustrate this. In planning an economy, an administrator has in mind the highest optimality as his goal. He sets up social arrangements and allocates natural resources accordingly in order to achieve this goal. It is possible that some members of society have to sacrifice a little in order that the total utility can be increased. Now, the planner, who considers the matter from the economic point of view, will think that this can be justified. In making social arrangements and in allocating available natural resources, all he has in mind is the increment of total utility for society, so his personal interest is not at stake. In this respect, we have to say that he is impartial. But, can we say that he has adopted the moral point of view? Rawls certainly will not think so because he believes there is a difference between justice and efficiency.

Footnotes

1. Kai Nielsen, "On Philosophic Method," International Philosophical Quarterly, Vol. XVI, No. 3 (Sept. 1976), pp.358-359.
2. These two methods are, of course, what we called earlier "the contractarian method" and "the method of reflective equilibrium." We shall discuss the equilibrium method later. In view of the fact that Rawls has been deeply influenced by Quine on foundational matters, we shall expect that he will have a holistic view about justification. This is exactly what we find. He says, "... justification is a mutual support of many considerations, of everything fitting together into one coherent view." See John Rawls, A Theory of Justice (Cambridge, Massachusetts, Harvard University Press, 1971), p.579.
3. Rawls, Ibid. p.121.
4. Ibid. p.121.
5. Ibid. p.44, and p.94. Of course, Rawls argues that the original position is being fair to everyone. However, since the contractors are self-interested and rational, their aim in the original position is to maximize their share of primary goods, we can describe them as prudential maximizers of primary goods.
6. Ibid. p.27.
7. Rawls' claim that the principle of utility is not a moral principle is ambiguous. Sometimes he can be understood as saying that it is not moral because it sanctions certain arrangements and actions contrary to our moral convictions; sometimes he seems to say that it is not moral because it is a principle of efficiency, and thus has nothing to do with morality. Footnote (6) clearly bears out the latter interpretation. Of course, the principle of efficiency can be immoral in the former sense when it requires arbitrary sacrifice of some people in society.

8. Rawls, op. cit., p.29.
9. Ibid. p.136.
10. Ibid. p.86.
11. Ibid. p.86.
12. Ibid. p.18.
13. Ronald Dworkin, "The Original Position," in Reading Rawls, ed. by Norman Daniels (New York: Basic Books Inc., 1975), pp.46-53.
14. Rawls, op. cit., p.12.
15. Ibid. p.4.
16. Ibid. p.10.
17. I borrow this example from G.B. Thomas' essay "On Choosing a Morality," Canadian Journal of Philosophy, Vol. V, No. 3 (1975), pp.357-74. He uses it for a different purpose from mine. He wants to show that moral principles are not the kind of things that can be chosen.
18. Rawls, op. cit., p.120
19. Aharon F. Kleinberger has done this in his article "The Social-contract Strategy for the Justification of Moral Principles," Journal of Moral Education, Vol. 5, No. 2 (1978), pp.107-126.
20. G.B. Thomas has done something like this in his article "On Choosing a Morality." op. cit. But his claim that the Rawlsian contractors will not understand what it is to be morally motivated to act seems to me to be false. The stipulation that the contractors are mutually disinterested only tells us that they will not be morally motivated to act, but it does not exclude the possibility that they can understand what it is like to be motivated to act because of moral considerations. For example, in the case of an egoist in ordinary life, who has never been morally motivated

to do anything and who doubts the value of adopting the moral point of view, it does not mean he does not understand what the moral point of view is. The fact is simply that he does not accept the moral point of view to guide his course of life.

21. Rawls, op. cit., p.136.

22. Ibid. p.127.

23. Ibid. p.148.

24. Ibid. p.136.

25. Ibid. p.153.

26. Ibid. p.154.

Chapter 4 The Method of Reflective Equilibrium

I. What is the Method of Reflective Equilibrium?

Aside from the contractarian method, there is another method Rawls uses to justify the two principles of justice he advocates. This is the method of reflective equilibrium, or, in short, the equilibrium method. Rawls first introduced this method in his article "Outline of a Decision Procedure for Ethics,"¹ and devoted the whole article to discussing it. In A Theory of Justice, Rawls' discussion of this method follows the general ideas outlined in the article, although it is presented in a more sketchy way. In our examination of this method, we shall make reference to both of these sources.

What is this method of reflective equilibrium? In order to answer this question, we need to have some idea of what is Rawls' conception of moral philosophy. He starts with the fact that under normal circumstances, each person beyond a certain age and possessing the requisite intellectual capacity develops a sense of justice. This moral capacity of ours exhibits itself in our skill in judging things to be just or unjust, right or wrong, and good or

bad, etc. He says:

Now one may think of moral philosophy at first (and I stress the provisional nature of this view) as the attempt to describe our moral capacity; or, in the present case, one may regard a theory of justice as describing our sense of justice. This enterprise is very difficult. For by such a description is not meant simply a list of the judgments on institutions and actions that we are prepared to render, accompanied with supporting reasons when these are offered. Rather, what is required is a formulation of a set of principles which, when conjoined to our beliefs and knowledge of the circumstances, would lead us to make these judgments with their supporting reasons were we to apply these principles conscientiously and intelligently. A conception of justice characterizes our moral sensibility when the everyday judgments we do make are in accordance with its principles. 2

With this conception of moral philosophy in mind, Rawls makes a comparison between the undertakings of moral philosophy and linguistics. The latter, as he conceives it, is the discipline whose "aim is to characterize the ability to recognize well-formed sentences by formulating clearly expressed principles which make the same discriminations as the native speakers."³ The aim of linguistics is then to describe and explain the sense of grammaticality the native speakers of a language have for the language.

Our sense of grammaticalness exhibits itself when we use our native language to utter well-formed sentences. So the data to be explained in linguistics are a set of well-formed sentences uttered by the native speakers of the language. By the same token, we exhibit our moral capacity when we make considered judgments in our daily life as to the matter concerning right, wrong, just, unjust, etc. The data to be explained in moral philosophy are a set of considered judgments we make concerning problems of "What is just and unjust," and "What is right and wrong," and so on. The considered judgments are "those judgments in which our moral capacities are most likely to be displayed without distortion."

After the characterization of "considered judgments," we come to the notion of "reflective equilibrium" itself. In view of the importance of this notion in Rawls' theory, we should quote Rawls' own discussion of it in full:

In searching for the most favored description of this situation we work from both ends. We begin by describing it so that it represents generally shared and preferably weak conditions. We then see if these conditions are strong enough to yield a significant set of principles. If not, we look for further premises equally reasonable. But if so, and these principles match our considered convictions of justice, then so far well and good. But presumably there will be discrepancies. In this case we have a choice. We can either modify the account of the initial situation or we can revise our existing judgments, for even the judgments we take provisionally as fixed points are liable

to revision. By going back and forth, sometimes altering the conditions of the contractual circumstances, at others withdrawing our judgments and conforming them to principle, I assume that eventually we shall find a description of the initial situation that both expresses reasonable conditions and yields principles which match our considered judgments duly pruned and adjusted. This state of affairs I refer to as reflective equilibrium. It is an equilibrium because at last our principles and judgments coincide; and it is reflective since we know to what principles our judgments conform and the premises of their derivation. 5

Now, we have all the essential notions involved in the equilibrium method. They are: (1) a set of considered judgments; (2) the device of the original position; (3) the explication of the considered judgments by a set of principles; (4) the notion of reflective equilibrium itself.

In "Outline of a Decision Procedure for Ethics," Rawls discusses another notion called "competent moral judges."

A person is qualified as a competent moral judge if he embodies certain characteristics, for example, he is reasonable and has sympathetic knowledge of a variety of human interests, etc. In the article, considered judgments are thought to be the judgments made by the competent moral judges. But in A Theory of Justice, Rawls seems to have abandoned this notion or at least he has not put so much emphasis on this idea of "moral experts." Instead, he asks

all of us, the readers, to join to exercise our moral capacities in judging the reasonableness of the conception of justice put forward in the book. Thus, he says, "We may suppose that everyone has in himself the whole form of a moral conception. So for the purpose of this book, the views of the reader and the author are the only ones that count." ⁶ In our discussion of the equilibrium method, we shall not examine the idea of "competent moral judges."

Now, let us explicate the four notions involved in the equilibrium method:

(1) Considered Judgments

First of all, not all the judgments we make daily concerning matters of morals can be used as data to check against the moral theories we advocate. Judgments are qualified as considered judgments and hence can be used to test a moral theory only if they embody certain features.

(a) It is necessary that the judgments be made under the circumstances that our own interests are not at stake; we do not stand to gain or loss by making the judgments. This condition is justified on the ground, Rawls argues, that fear and partiality are recognized obstacles in the determination of justice.

(b) It is required that before making the judgments, all the relevant facts concerning the case be known to us. This requirement can be justified on the ground that correct moral judgments can hardly be obtained if certain facts relevant to the case are hidden from us.

(c) It is essential that we must have great confidence in the judgments we make. This characteristic is called "certitude" by Rawls in contrast to "certainty," which is a logical relation between a proposition, or a theory, and its evidence. This requirement is essential because there can hardly be any point to say that we can use a set of judgments to test a theory if the judgments seem to us to be wrong or a little uncertain.

(d) It is necessary that the judgments be intuitive with respect to ethical principles, that is, we must not consciously invoke any moral theory to help us in making the judgments. This requirement is crucial in view of the aim of Rawls's theory. Its aim is to show us that the two principles of justice of his theory represent the most plausible conception of justice, and the reason for this is that they match our considered judgments. If considered judgments are made by our conscious application

of certain ethical principles, they will no longer enjoy the neutral status they are supposed to embody. As a matter of fact, the whole method of reflective equilibrium will be threatened with circularity.

A final note as to the notion of considered judgements is that they are not restricted to the lower-level judgments about particular cases of moral issues. Rawls claims, "People have considered judgments at all levels of generality from those about particular situations and institutions up through broad standards and first principles to formal and abstract conditions on moral conceptions."⁷

(2) The Original Position

The notion of the original position is mainly used in the contractarian method as the initial status quo. From there, the contractors are to reach an agreement. It contains all the premises used in the derivation of the principles of justice. But the design of the original position needs help from our considered judgments. The reason we eventually accept a certain interpretation of the initial contractual situation as the most plausible one is that the principles derived from it match our considered judgments best. This shows that there is a close relation⁰

between the contractarian method and the equilibrium method. The relationship between the two methods is complicated. For us, it suffices to note that the reasonableness of the device of the original position depends upon whether the principles derived from it agree with the judgments we make on the basis of our sense of justice.

(3) Explication of our Considered Judgments

One of the principal tasks of an ethical theory is to derive a set of principles which can be used in determining and in explaining the crucial problems of "what is right and wrong," and "what is just and unjust," etc., when the need of making such kind of judgments arises. This is certainly one of Rawls' chief aims in A Theory of Justice. This task is mainly the job of the contractarian method. In the book, the principles of justice are derived from a social contract.

Once a class of considered judgments has been selected in an intuitive way, and a set of principles has been derived from a social contract, what remains for us to do is to see whether they match. When the principles and our considered judgments match, Rawls conceives the relation the principles bear to the considered judgments

to be one of explication. Our considered judgments simply state what is right and wrong, just and unjust, without providing reasons. They need to be explicated, in the sense that reasons must be offered as to why they are right and just. A set of moral principles serves this purpose,

...an explication of these judgments is defined to be a set of principles, such that, if any competent man were to apply them intelligently and consistently to the same cases under review, his judgments, made systematically non-intuitive by the explicit and conscious use of the principles, would be, nevertheless, identical, case by case, with the considered judgments of the group of competent judges. 8

Rawls then cautions us that we must not confuse certain things with an explication of our considered judgments. First, an explication is not an analysis of the meaning of the ethical terms used in the judgments constituting its range. Secondly, an explication is not concerned with what people intend to express when they make moral judgments, for example, an explication is not concerned with the problem of whether people are only expressing their emotions when they make moral judgments. Thirdly, an explication is not a theory about the cause of the considered judgments, that is, it is not concerned with what motivates people to make considered judgments. In other words, it has nothing to do with people's psychology. Finally, an expli-

cation is unsatisfactory if we can show that the principles employed to explicate our considered judgments fail to yield certain judgments within its range. Conversely, an explication is satisfactory if we can show that the principles and our considered judgments match completely. That is to say, we can derive a set of considered judgments from the principles together with certain statements describing the relevant concrete situations.

(4) Reflective Equilibrium

Finally, we come to the notion of reflective equilibrium itself. According to the provisional aim of moral philosophy as conceived by Rawls, we can say that his theory (Justice as Fairness) involves the hypothesis that the principles which would be agreed to in the original position are identical with those that match our considered judgments. But, of course, in the event of constructing a moral theory or any theory for that matter, we cannot expect to be so fortunate that in our very first attempt, we shall be able to arrive at the principles which match our considered judgments completely. Presumably, there will be some gaps. The equilibrium method instructs us to work both ways when there is a discrepancy between the

principles derived from the original position and our considered judgments. Sometimes we stick to our considered judgments when we have great confidence in them and change the conditions imposed on the original position; and sometimes we revise our considered judgments to make them conform to the principles. By shuttling back and forth, we can hopefully come to a complete match between our considered judgments and the principles derived from the original position. This state of affairs Rawls calls "reflective equilibrium." It is an equilibrium because the principles and our considered judgments are in perfect match; and it is reflective because we know to what principles our judgments conform, just as we know the premises of their derivation.

The equilibrium method for the justification of moral principles can be summarized as follows: a moral theory is not different from any other scientific theories. It starts with the problem of how to explain our moral capacity which exhibits itself by the ability we have in making judgments as to what is right and wrong, just and unjust, and so on. First, we select a set of considered judgments as our data to be explained. Then we select a set of con-

ditions as our starting point and from it we derive a set of principles which can be used to explicate and explain the considered judgments we make intuitively. In the course of constructing the theory, we constantly shuttle back and forth to revise our considered judgments and the principles in order to make them a perfect match. When this state of affairs is achieved, our work is completed. In this state of reflective equilibrium, the considered judgments are explained by the principles we finally arrive at; the principles and the theory are justified or confirmed by the considered judgments. Similarly, in other scientific theories, the observational data are explained or explicated by a scientific hypothesis, and the hypothesis itself in turn is confirmed or justified by the observational data. Rawls himself condones this analogy between scientific theories and moral theory. He says: "I wish to stress that a theory of justice is precisely that, namely, a theory. It is a theory of the moral sentiments...or, more specifically, our sense of justice."

II. The Method of Reflective Equilibrium and Relativism

In doing moral philosophy, there are two tasks a theory must accomplish: first, a theory must prove to be a moral theory, that is to say, it must provide grounds and reasons for us to believe that the theory has characterized a moral point of view; and second, it must show that the theory is superior to other moral theories. Accordingly, the word "justification" has two meanings for a moral theory. First, when it is claimed that a certain method constitutes a justification of certain principles of that theory, what we mean by this is that by employing the method, we are justified in claiming that the principles derived are moral principles, that is, they have something to do with morality or are acceptable from a moral point of view. Second, "justification" can also be used in connection with ranking different moral theories, that is, to judge that a moral theory as a good or a bad one, or that it is better than other moral theories. In this case, when we say that a certain moral theory is justified, we mean that the theory in question is superior to other moral theories. Any moral theory must have some way of coping

with the first sense of the word "justification" in order to be qualified as a moral theory. In connection with the discussion of the contractarian method, we have argued at length that the method cannot successfully deal with this issue. It cannot prove that the principles reached by the contractors are necessarily principles of justice, and we have no way of knowing why the principles agreed to by the contractors have anything to do with morality. In other words, the contractarian method does not have moral force. This task is for the equilibrium method to achieve. This method demands that the principles reached by the contractors match our considered moral judgments. In this way, the relevance of the principles to morality can be established. The way it establishes the relevance is as follows: (1) The conditions embodied in the original position contain certain special features peculiar to moral theories. This is so because in designing the original position, we come to those conditions with the help of our considered judgments. We examine our considered judgments constantly to see if the conditions imposed on the original position will yield principles which will match our considered moral convictions. (2) The principles derived from

the original position must match our considered moral judgments as to what is right and wrong, just and unjust, etc.

In criticizing the equilibrium method, we shall concentrate on the problem of whether the method can be used successfully in ranking different moral theories. Rawls apparently believes it can, but our answer to this question is negative.

In the last paragraph of A Theory of Justice, Rawls makes a final attempt to convince us about the plausibility of the device of the original position:

Each aspect of the original position can be given a supporting explanation...Thus what we are doing is to combine into one conception the totality of conditions that we are ready upon due reflection to recognize as reasonable in our conduct with regard to one another (§ 4). Once we grasp this conception, we can at any time look at the social world from the required point of view...It enables us to be impartial, ...Thus to see our place in society from the perspective of this position is to see it sub specie aeternitatis: it is to regard the human situation not only from all social but also from all temporal points of view. 10

Since the original position represents such a point of view, correlatively, the principles derived from it will be universally valid for all people from different cultures and historical periods. This is a rather strong claim. It is tantamount to saying that both ethical and meta-ethical

11
relativism are wrong. Ethical relativism is wrong because the original position represents the moral point of view, and the principles derived from it are valid for all people living in different cultures and historical periods. Meta-ethical relativism is wrong because there is a sound procedure with which we can validate a moral theory. In other words, Rawls thinks he has solved the problem of how to rank various moral theories. But is this really so? Has relativism really been refuted? In connection with the discussion of the contractarian method, we have pointed out that the method cannot successfully deal with the problem of ranking different ethical theories. The argument that a set of principles would be chosen by the contractors in the original position does not constitute a justification for the claim that the set of principles is superior to other principles. This is so because the reason the contractors will choose a certain set of principles over other principles is because the conditions imposed on the original position make them do so. If we interpreted the initial contractual situation differently, a different set of principles would be preferred. This proves that the burden of the justification is carried by the condi-

tions in the original position rather than the notion of a social contract. We shall examine presently whether the equilibrium method can successfully carry out this mission. If it can, then the claim that the two principles of justice advocated by Rawls have universal validity will be justified.

The equilibrium method states that in testing the validity of moral principles, what we must do is to check them with "our" considered moral judgments about what is good and bad, just and unjust, right and wrong. If the principles match those considered judgments of ours, this fact constitutes a justification of the principles.

Steven Lukes, in his article "Relativism: Cognitive and Moral,"¹² expresses some uneasiness about the appeal to "our" and "we" in justifying moral principles. He points out that it is a fact that we obviously do not agree on very many important moral issues; when this occurs, the appeal to our considered judgments for the justification of ethical principles can run into trouble. Even if we restrict ourselves to the contemporary Western society, we can still find such different groups of people as "ultra-conservatives, clerical authoritarians, empire loyalists, fascists, racial separatists, Saint-Simonian

technocrats, individualist liberal, anarchists, radical
egalitarians,..."¹³ It is plausible to say that we can
hardly expect them to have the same considered judgments
on every moral problems. Indeed, it is very likely that
they will disagree on each and every important moral issue. 3
When disagreements among them occur, the equilibrium me-
thod does not give us any hint as to what we can appeal
to in settling the disagreements. In Rawls' case, the
particular group or audience he appeals to are people who
represent the liberal-social democratic tradition. They
are people who believe a world will be more just if,

there would be less inequality of wealth, no
inherited privilege, no educational discrimina-
tion 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,...¹⁴

These are the beliefs obviously embedded in the ori-
ginal position. They clearly reflect the values of the
liberal, individualistic and democratic tradition of cer-
tain modern Western people. As we have said before, we
can hardly expect someone who is a fascist to agree with
all these values. When disagreements between a Rawlsian
liberal and a fascist occur, Rawls gives no reason as to
why we have to accept the values embodied in the considered

judgments of the liberal democrats rather than the values accepted by the fascist. It is certainly of no use if Rawls simply states that it is because the considered judgments made by the liberal democrat are the right ones, the ones any reasonable men should accept. There is no way for us to know that this is the case. If we take the fact seriously that there are such divergent groups of people in modern Western society, ranging from fascists to anarchists, each holding his own beliefs about what is right and wrong, just and unjust, then we can hardly expect that the appeal to our considered judgments will accomplish the kind of mission Rawls believes it can, that is, to prove that a set of principles of justice is universally valid. This is so for the following simple reason. Anybody disagreeing with the values of the liberal tradition can ask why he has to regard those values as superior to his own and use them to check against the moral principles of a proposed theory. Why can he not, as a fascist, in testing the plausibility of a moral theory, appeal to his considered judgments, which are presumably radically different from the ones accepted by the liberal?

Rawls is aware of the difficulties involved in the use of our considered judgments for validating moral principles. Thus he remarks:

second, there is the more specific objection to the argument I have presented that it depends upon a particular list of conceptions of justice between which the parties in the original position are to choose, and it assumes not only an agreement among persons in their considered judgments, but also in what they regard as reasonable conditions to impose on the choice of first principles. 15

But curiously enough, after Rawls has anticipated such an objection, he has not given any answer to it. It is a rather strong and valid objection to the equilibrium method. He once remarks:

...justification proceeds from what all parties to the discussion hold in common. Ideally, to justify a conception of justice to someone is to give him a proof of its principles from premises that we both accept, these principles having in turn consequences that match our considered judgments. 16

But this remark is of no help. What it says is that to justify a theory of justice or a moral theory, we must start from some premises accepted by all parties to the discussion, and it is also required that their considered judgments must be in agreement, otherwise the relation between the premises and the conclusion are simply that of a proof. But proof is not justification. The objection to Rawls' method of appealing to consensus is raised

precisely because the opponents of the equilibrium method think that the way Rawls derives his principles constitutes only a proof but not a justification. This is so because: (1) we do not have consensus as to what conditions are reasonable to be imposed on the initial choice situation, and (2) people's considered judgments do differ considerably as to what is right and wrong, just and unjust. The following remark by Rawls is a little evasive: "It may be said that the agreement in considered judgments is constantly changing and varies between one society, or part thereof, and another."¹⁷ As Kai Nielsen points out: "Why say 'it may be said'? Is it not, if we give a reasonable interpretation to 'constantly', plainly so?" It would surely seem to be at least."¹⁸

In view of the above analysis, what Rawls has to do is to show us that there is a set of considered judgments, which is universally valid for each and every person from different cultures and historical periods. Or, he should at least argue that the set of considered judgments he (and the readers) appeals to has such universal validity. But this is what Rawls has not shown us. He has done nothing to show us why the considered judgments he appeals

to, which represent the considered judgments of certain modern Western liberals, are valid for all people. Nor has he shown us why in testing a moral theory, these considered judgments should be used instead of some considered judgments held by people who believe in Fascism.

The success or failure of any method must be judged in view of what it attempts to achieve. The equilibrium method is designed to achieve a very ambitious goal of proving that the principles which can be shown to match our considered judgments are universally valid. But we have pointed out that because of the very nature of our considered judgments, this goal cannot be obtained. The sociological fact that different people have different considered judgments will defeat the ambition of validating moral principles as universally valid by the equilibrium method. In natural science, if we follow the account offered by the logical positivists, there is a set of observational sentences which is true independent of any theory.¹⁹ But in the realm of morality, such a set of judgments is not available to us. It is puzzling for me that Rawls, who is so deeply influenced on foundational matters by Quine, can take such a positivistic stand in

believing that there is a set of considered judgments which can play the role of what Popper calls "basic statements" in testing the validity of moral theories.

It seems to me here that the fact that Rawls must appeal to particular considered judgments which are peculiarly liberal and democratic, such as societies which are intolerant of heresy are unjust, societies committed to hereditary hierarchies are unjust, conclusively shows that relativism has not been refuted. When questioned by Plato or Aristotle, who subscribed to the view that natural inequalities should be reflected in our social institutions, as to why modern liberal democratic values should be accepted, Rawls can do nothing except repeat his conviction that these considered judgments are more reasonable. But convictions are not arguments.

What we have done so far is simply point out that the appeal to our considered judgments in validating moral theories will run into trouble because of the sociological fact that different people have different considered moral convictions as to what is right and wrong, just and unjust, etc. Now we want to examine the equilibrium method from a different angle. The way we shall proceed is as follows:

Given the facts that there are different conceptions of justice, for example, utilitarianism, perfectionism, etc., and that different people hold different conceptions of justice, is it possible for all people to have the same considered moral judgments or, at least, a shared core group of such judgments? It is important to ask this question because for the equilibrium method to be effective we must assume that it is at least possible that there is a set of considered judgments which is true independent of all theories, that is, the truth of these judgments can be established without appealing to any principles of justice or morality which pertain to a particular conception of justice.

Now, it is a fact that there are rival theories of justice and different people hold different conceptions of justice. Each conception of justice, according to the equilibrium method, entails the set of considered judgments of that theory. Let us recall briefly how the equilibrium method proceeds. The provisional aim of moral philosophy is an attempt, according to Rawls, to explicate our moral sentiments which are exhibited by the capacity we have in making considered moral judgments. Let us imagine a person who has been making judgments of the type all his life,

but has never had a clear conception of what sort of moral theory he has subscribed to. Now he intends to make it clear to himself, and he wants to do a little philosophizing. Let us further assume that he is utilitarian-oriented because the judgments he has been making exhibit that they very often conform to the principle of utility. For instance, he firmly believes that it is right for someone to be made to sacrifice in certain situations in order that the group the person belongs to might gain more.

He starts to formulate a set of abstract principles to explicate the more concrete judgments he makes. There are two possible cases we must consider. The first possibility is as follows: suppose after great deliberation which involves comparing various normative ethical theories, he comes up with a sort of utilitarian theory. Now, he is in a state of equilibrium. That is to say, the set of principles of justice he has arrived at corresponds to the moral considered judgments he has been making intuitively, and his considered judgments are explicated by the principle of utility. If he uses the principle of utility, together with the concrete situations of a particular case, he will be able to derive the judgments which conform to the judgments he makes intuitively without the invocation

of the principle of utility. Of course, it is highly unlikely that he can hit upon the principle of utility the first time he proposes a certain set of principles to himself. As the equilibrium method suggests, it will take much painful experience of shuttling back and forth between the principles and the considered judgments in order to achieve the state of reflective equilibrium. He must also take other available ethical theories into consideration. It involves many trials and errors. But the important thing to note is that he has finally reached this state of reflective equilibrium.

The second possibility is that when he tries to formulate a set of abstract principles, he is struck by the reasonableness and plausibility of Rawls' two principles of justice. This will lead him to find out that there are many discrepancies between his considered judgments and the two principles he presently thinks acceptable. Let us suppose that the reasonableness of the two principles has such an overwhelming power over him that he is willing to give up many considered judgments he has accepted as reasonable so far because they are in conflict with the two principles he now considers plausible. Again, after

great deliberation of going back and forth between his considered judgments and the principles, he reaches a state of reflective equilibrium. The difference in this case is that he gives up all those considered judgments which are in conflict with the two principles, even if he has previously accepted them as reasonable. All the considered judgments he accepts now are in conformity with the two principles of justice. We can rightfully characterize him as having undergone a conversion, that is, he has changed his moral position from a utilitarian one to a Rawlsian one, even if he is not aware of the fact that he had held a utilitarian position before, and that he did not explicitly subscribe to the principle of utility. Now, suppose a utilitarian moralist comes along and tells him that the theory he now subscribes to is wrong, and that the considered judgments he has abandoned in the course of formulating the theory were the right ones, what kind of response can we expect from him? Most likely, we shall see him vigorously defending this new theory he presently accepts as reasonable, and remaining truthful to the new set of considered judgments which conforms to the two principles of justice. In this case, there seems to be nothing the

utilitarian moralist can do. He is not likely to give up, after such great deliberation, his theory and the corresponding judgments.

To make it possible for us to compare and evaluate the principles underlying the different theories, there must be either a universal standard which can be used to assess theories or a set of judgments, which is known to be true independent of any theory, so that all these different theories accept them as their theorems.²⁰ But, unfortunately, as we have shown before, we can have neither of them. The real difficulty of ethical dispute is that it is no use when two persons, A and B, disagree on something concerning morals, if A says to B, "Your judgments are not well-considered, you have not gone through great deliberation," while B can say exactly the same thing to A, and there is nothing they can appeal to, which is acceptable to both and will settle their disagreement. If they are philosophers, they will finally formulate fundamentally different principles which will explicate their own considered judgments respectively. If they start from the disagreements about ethical principles, what they can appeal to is their considered judgments in order to justify

their own view, but they will find they disagree on them too, because considered judgments are entailed by the corresponding principles. The equilibrium method seems to make us move in a circle. The principles and the corresponding judgments form a consistent set.

We must ask why this is the case. A set S of statements, is inconsistent if, and only if, for any statement P , P can be derived from S , for example, the set " $\{P \supset Q, P, Q\}$ " is inconsistent because by the theorem " $\{A, \sim A\} \vdash B$ for any B ," and Modus Ponens, we will be able to derive r from the set while r is any arbitrary statement. When we say that two theories are rival theories, for example utilitarianism and Justice as Fairness, what we mean is that the set consists of the principle of utility and the two Rawlsian principles of justice is inconsistent. Hence, for any statement P , P is derivable from the set. The ground for me to say that the set consisting of the principle of utility and the two principles of justice advocated by Rawls is inconsistent, is that, from the principle of utility we will be able to derive the statement that for the sake of total utility, sometimes we must sacrifice the interests of certain members of a group, but from Rawls'

two principles, we will derive the negation of this statement, that is, it is never the case that we must sacrifice the interests of certain members of a group in order to increase the total utility. The former will be held as a fixed considered judgments and irreversible by a utilitarian, but it is wrong for a Rawlsian moralist. Even though the state of reflective equilibrium is a state reached after one has weighed various conceptions of justice, it is not unique, People holding different theories will have their own equilibria. The state of reflective equilibrium for a utilitarian is a state of complete match between the principle he holds and his considered judgments; for a Rawlsian moralist, it is the state of complete match between the two principles of justice and his considered judgments. When a Rawlsian moralist measures some of his most fixed considered judgments against the principle of utility, he is bound to discover some discrepancies for the very reason that utilitarianism and Justice as Fairness are rival theories. Some of the statements derivable from the principle of utility are in contradiction with the considered judgments a Rawlsian moralist subscribes to. In case a discrepancy occurs, the Rawlsian moralist

has a choice of either repudiating the principle of utility, or revising his considered judgments. If he does the former, he remains a Rawlsian, and, accordingly, he will reject those considered judgments entailed by the principle of utility, which are in conflict with the two principles of justice; on the other hand, if he does the latter, he will cease to be a Rawlsian and is converted into a utilitarian. ²¹ It seems we have to conclude that it is impossible to have a set of fixed considered judgments in reflective equilibrium whose truth is independent of all theories. When we say that the principles and their corresponding judgments form a consistent set, what we mean is that between rival theories, there always seems to be some incommensurability involved.

In view of the above analysis, it seems to me, Rawls' attempt to refute ethical relativism by employing the equilibrium method is a failure. Although one of the conditions of the original position states that the principles proposed for acceptance must satisfy the condition of publicity and even when the contractarian approach does satisfy this condition, it still does not follow from this that Justice as Fairness will be accepted universally. The utilitarians,

whose considered judgments are in some measure different from those of a Rawlsian moralist, can provide a different interpretation of the initial contractual situation and derive different principles of justice. It will also satisfy the condition of publicity. By appealing to our considered judgments for the justification of a theory, the most we can achieve is to come up with a conception of justice so that its principles match the considered judgments we currently hold. It is still a long way to go for Rawls to justifiably claim that he has constructed a theory which will universally be accepted by all people with divergent considered judgments. Perhaps there is no such position in morality from which we can view things sub specie aeternitatis.

III. The Method of Reflective Equilibrium and Intuitionism

Rawls considers intuitionism one of the main rivals to his theory. In connection with the discussion of the problems of the priority of liberty and of how to order various first principles, he examines it extensively. He denies that his theory is intuitionistic despite the fact that

he acknowledges in several places that in the course of constructing his theory, he cannot avoid the appeal to intuition entirely.

In his article "Rawls's theory of Justice," Hare has argued that Rawls is a subjectivist, and he further claims that intuitionism is nearly always a form of disguised subjectivism and that Rawls is an intuitionist.²²

In this section, I intend to argue for three things: first, Rawls is not a subjectivist in the sense Hare describes; second, even though Hare is right in labelling Rawls as an intuitionist, the reasons he gives for this claim are wrong; third, I want to show in what sense and for what reason Rawls is an intuitionist.

Rawls characterizes intuitionism in the following way:

Intuitionist theories, then, have two features: first, they consist of a plurality of first principles which may conflict to give contrary directives in particular types of cases; and second, they include no explicit method, no priority rules; for weighing these principles against one another.²³

According to this definition, we can make the following three points about intuitionism: first, intuitionism, as characterized by Rawls, is a meta-ethical thesis rather

than a substantive moral theory, since it has not told us anything about the content of theory being labelled as intuitionistic, when we claim that a certain moral theory is intuitionistic, we are making a meta-ethical statement. Second, an intuitionistic theory must contain more than one substantive moral principles, and there is a possibility that the different principles will give different directives in a certain situation; this rules out a theory like utilitarianism as being intuitionistic, since it consists of only one ultimate moral principle. Third, when conflicts between the first principles occur, there is no higher standard to which we can appeal in order to settle the conflicts; the only thing we can resort to in this situation is our intuition.

But intuitionism construed in this way is certainly too restricted, and it does not square with the traditional sense of the word. Intuitionism in the traditional sense includes certain moral-epistemological theses, for instance, those concerning the self-evidence and the necessity of moral principles. The most fundamental one of these theses states that the way to grasp the validity of moral principles is by intuition. Rawls' definition of intuitionism

will rule out philosopher like G.E. Moore as intuitionist because he is a utilitarian; and being a utilitarian, he accepts only one principle, namely, the principle of utility, as the ultimate principle. Any other principles, for example, equal distribution of goods, maximum liberty for everyone, will be subordinate to the principle of utility. But we would certainly like to consider Moore an intuitionist, and so does Rawls.²⁴ Rawls does not seem to realize that for a moral theory, even if it does not contain plurality of first principles, the problem of whether intuition is the final resort for the validity of moral principles can still arise. Hare believes that, the right name for the kind of intuitionism Rawls has in mind would be "pluralistic intuitionism." But he continues, "There can also be another, non-pluralistic kind of intuitionist -- one who intuit the validity of a single principle or ordered system of them, or of a single method, and erect his entire structure of moral thought on this."²⁵ Thus, we shall conceive intuitionism in a more general way than Rawls does. We believe that the definition we are going to give not only squares with the traditional sense of the word better but also captures the essential of the definition Rawls gives to it.

It will be conceived as a meta-ethical position which states that "in the process of validating an ethical theory, or in the event of applying moral principles to concrete situations, we inevitably have to rely on our intuition."

Rawls clearly denies that his theory is intuitionistic in either sense of the word. It is not intuitionistic in his own sense of the word because, despite the fact that his theory contains two ultimate principles of justice, it has solved the problem of their priority. The principles are ordered lexically (whether Rawls has successfully proved this is a matter of dispute which I shall not discuss here). It is not intuitionistic in our sense of the word because, despite Rawls' own acknowledgement that he relies on intuition heavily throughout the book, it is possible to avoid the resort to intuition entirely in constructing a theory of justice. Thus he says:

The argument aims eventually to be strictly deductive. To be sure, the persons in the original position have a certain psychology, since various assumptions are made about their beliefs and interests. These assumptions appear along with other premises in the description of this initial situation. But clearly arguments from such premises can be fully deductive, as theories in politics and economics attest. We should strive for a kind of moral geometry with all the rigor which this name connotes. 26

As stated earlier, Hare claims that Rawls is advocating a kind of subjectivism, and he also argues that Rawls is an intuitionist. He makes both of these claims on the basis that Rawls uses the method of reflective equilibrium for validating moral principles. We believe that Hare's first claim that Rawls is a subjectivist is wrong, and that even though he is right in labelling Rawls as an intuitionist, the reasons he gives for the claim are wrong.

Hare argues that Rawls' theory is subjectivistic on the basis that Rawls appeals to our considered judgments for the validation of moral principles. The appeal to our considered judgments in moral reasoning, according to Hare, makes the answer to the question "Am I right in what I say about moral questions?" depend on the answer to the question "Do you, the reader and I agree on what we say?"²⁷

In addition to this, Hare claims that a subjectivist is a person who holds a meta-ethical position which states that the final court of decision for the validity of moral principles is people's views on the principles. The subjectivist further believes, that when there is a conflict between the principles and our considered judgments, it is always the principles that must be repudiated. He says:

The element of subjectivism enters only when a philosopher claims that he can 'check' his theory against his and other people's views, so that a disagreement between the theory and the views tells against the theory. To speak like this (as Rawls does constantly throughout the book) is to make the truth of the theory depend on agreement with people's opinions. 28

It is true that the method of reflective equilibrium requires us to check the principles of a moral theory against our considered judgments. But it is a big mistake to say that the method instructs that if there is a conflict between the proposed principles and our considered judgments, it is always the principles that must go. To say so is to forget the most important feature of the equilibrium method. Let us recall briefly what the equilibrium method teaches us. Rawls acknowledges that in the course of constructing a moral theory by using our considered judgments as measures for validating moral principles, most likely there will be discrepancies between the principles of the theory and our considered judgments. When the discrepancies occur, it is not necessary that we reject one or the other. Sometimes, we have great confidence in our considered judgments, and in those cases, we reject the proposed principles which do not accord

with them; and sometimes we have to reject our considered judgments in order to accomodate our theory. It is a method which works from both ends. We must shuttle back and forth between our considered judgments on the one hand, and the principles of a theory on the other. And this process of going back and forth must carry on to the point when we reach a complete match between the two. It seems that the kind of element of subjectivism described by Hare does not really play a role here. Rawls will never say that when there is a disagreement between the theory and our considered judgments, it is always an evidence to tell against the principles.²⁹ Rawls has never been a subjectivist in Hare's sense. It seems to me that Hare is simply attacking a strawman.

Secondly, Hare argues that Rawls is an intuitionist. Hare's conception of intuitionism is more or less similar to the one we subscribe to, namely, a moral theory is intuitionistic if and only if it asserts that in the course of validating an ethical theory, or of applying moral principles in concrete situation, we inevitably have to rely on our intuition. Hare makes his criticism on the basis that Rawls must appeal to intuition in two places in con-

structing his theory. He says:

Thus intuition can enter at two points (and in Rawls' case enters at both; compare 121/7-15). It enters in the choice of the conditions to which the chooser is to be subject; and it enters to determine what he will choose in cases where the conditions, as made explicit, do not determine this. 30

Let us consider the second point first. What Hare is talking about here concerns the problem of the derivation of moral principles. Rawls admits that, as a matter of fact, he has relied heavily on intuition in deriving the two principles of justice from the original position. But this does not imply that theoretically intuition cannot totally be avoided. We only have to recall his notion of moral geometry. He believes that it is quite possible that in ethics, we can derive the principles in the strictly deductive manner from the premises as ~~it~~ has been done in theories of politics and economics. Perhaps Rawls' derivation of the two principles of justice from the original position is not a satisfactory one. But an unsatisfactory deduction is one thing, and intuition is quite another. The crucial point is whether or not in moral reasoning we inevitably have to rely on our intuition to derive the principles. I believe that Rawls' use of the contractarian

model has satisfactorily shown us that in moral reasoning, strict deduction is not an unattainable goal notwithstanding the many difficulties it faces. One of the most important aims of the contractarian approach to moral philosophy is to provide a way of replacing intuitive or controversial moral judgments by judgments of rational prudence. By transforming the problem of moral choice to the problem of rational choice, the conditions the choosers are subject to will determine what principles they will choose.

As to the first point, even if we have to rely on our intuition in the choice of the conditions to which the contractors are to subject, we have to keep in mind that our intuition does not have the final say here, i.e., our considered judgments will sometimes have to be abandoned in view of the plausibility of the principles we arrive at. Intuition is by no means the final court of appeal. Whether the conditions are reasonable must depend on the principles we finally arrive at. Further, Rawls claims that "each aspect of the original position can be given a supporting explanation."³¹ According to the equilibrium method, the conditions we choose as the starting point for the derivation of principles are always tentative. They are subject to revision in view of the principles derived

from them. Whether the conditions are reasonable will depend on two things: (1) our intuition; (2) the principles derived. But the important thing here is that intuition is not the only thing we can and must appeal to. The relationship between the principles and our considered judgments is one of constant mutual criticism. Sometimes we use the principles derived as our point of reference to criticize and reverse our intuition, and sometimes it is the other way round. At this stage, our intuition does not have the unique authority alluded to by Hare. The process of going back and forth between the principles and our considered judgments is in constant play until we reach the state of reflective equilibrium.

The way Rawls proceeds in constructing his theory is probably as follows: first, he has a firm belief in the two principles of justice (they may be somewhat different from the final version he gives), then he stipulates a set of conditions which will yield these two principles, and finally he checks them with his considered judgments and sees if they match. Of course, there must be many revisions involved in both ends before finally reaching the state of reflective equilibrium, that is, the complete match.

between the two principles and his considered judgments. If this conjecture is correct, the way Rawls proceeds is not really different from the way natural scientists do in formulating their theories. When scientists first propose certain hypothesis, they must also rely on intuition. But we are never bothered by the fact that scientists rely on their intuitions in the course of theory-construction. This is because we believe there is a way they can test the validity of their theories. Similarly, in proposing a moral theory, we should not be bothered by the fact that we have to start from some intuition as long as we have some way of validating our theories. As a matter of fact, aside from our intuition, where can we start? This is true of both scientific theories and ethics. "Justification is," as Rawls remarks, "a matter of the mutual support of many considerations, of everything fitting together into one coherent view."³²

Finally, we come to the question: "Why do we consider Rawls an intuitionist?" In order to answer this question, let us recall once more our definition of the term "intuitionism." Intuitionism is a meta-ethical position which states that in the course of validating an ethical

theory, or in the event of applying moral principles to concrete situations, we inevitably have to rely on our intuition. Thus, in order to answer our initial question affirmatively, we have to point out where Rawls inevitably has to rely on intuition in validating the principles of justice. It seems to me there is one place he cannot avoid the appeal to intuition when he uses the equilibrium method. The appeal to intuition must take place when there is a conflict between the proposed principles and his considered judgments. When conflicts occur between his considered judgments and the principles, there is no higher court he can appeal to aside from his intuition for favoring the one or the other. In such cases, his intuition must decide whether he should give up his considered judgments or revise the proposed principles. So, despite Rawls' denial that he is an intuitionist, it seems to me that he should be considered as one in the traditional sense of the word.

Footnotes

1. John Rawls, "Outline of a Decision-Procedure for Ethics," The Philosophical Review, 60 (1951), 177-97.
2. John Rawls, A Theory of Justice, (Cambridge, Massachusetts, Harvard University Press, 1971), p.46.
3. Ibid. p.47. Thomas Nagel and R.M. Hare both believe that this analogy is false for the reason that the intuitions of native speakers are decisive as regards grammar. People's linguistic intuitions are indeed, in the end, authoritative for what is correct in their language. But whatever we ordinarily agree in praising is not necessarily right and what we ordinarily agree in condemning is not always wrong. See Nagel's article "Rawls on Justice" and Hare's article "Rawls' Theory of Justice," in Reading Rawls, ed. by Norman Daniels. (New York: Basic Books Inc., 1975), p.2 and p.86.
4. Rawls, A Theory of Justice, p.47.
5. Ibid. p.20.
6. Ibid. p.50.
7. John Rawls, "The Independence of Moral Theory," Presidential Address to the American Philosophical Association, Eastern Division, 1974, in Proceedings and Addresses of the American Philosophical Association, 48 (1975), p.8.
8. Rawls, "Outline of a Decision Procedure for Ethics," p.184.
9. Ibid. pp. 50-51.
10. Rawls, A Theory of Justice, p.587.
11. By ethical relativism I mean the normative thesis that what is right or good for one person or society is not necessarily right or good for another person or society even though the situations involved are similar. By meta-ethical relativism I mean the thesis that there is no sound procedure for justifying the validity of a moral theory. For the discussion of various forms of relativism, see Kai Nielsen's article "On the Diversity of Moral Beliefs," Cultural Hermeneutics 2 (1974), pp.281-303.

12. Steven Lukes, "Relativism: Cognitive and Moral," The Aristotelian Society Supplementary Vol XLVIII (1974), pp. 165-89.
13. Ibid. pp.182-183.
14. Ibid. p.182.
15. Rawls, A Theory of Justice, p.580.
16. Ibid. pp.580-581.
17. Ibid. p.580.
18. Ibid. p.41.
19. Kuhn and Feyerabend will of course dispute the availability of such a set of sentences even in natural sciences. Their arguments depend on the contention that there is no real distinction between observational terms and theoretical terms. All observational terms are theory-laden.
20. By "universal standard" I mean a set of criteria which can be used in assessing the plausibility of competing theories. Philosophers of science argue that such a set of criteria exists in the scientific enterprise. Even a relativist like Kuhn, who advocates the incommensurability thesis, states that accuracy, scope, simplicity, fruitfulness, and the like are universally accepted by scientists as valid criteria in making theory choice. It can be questioned whether universal acceptability constitutes validity. The case of normative ethical theory is even more difficult. In normative theory, the notions of "rationality," "good," "better than," etc., themselves are the subjects of investigation. See Thomas Kuhn, "Reflections on my Critics," in Criticism and the Growth of Knowledge, ed. by Imre Lakatos and Alan Musgrave (Cambridge: Cambridge University Press, 1970), 231-278. For the notion of basic statement, see Karl Popper, Conjectures and Refutations (New York: Harper & Row, 1963), especially essays 10 and 11, from pp. 215-280.

21. Of course it is extremely difficult to describe this sort of moral conversion. It is hard to pinpoint precisely when he ceases to be a Rawlsian and becomes a utilitarian. The situation is similar to what Kuhn calls "paradigm-shift" in a period of scientific revolution. For Kuhn's notion of paradigm-shift, see his book The Structure of Scientific Revolutions (Chicago: The University of Chicago Press, 1962), Especially VI-XII.
22. R.M. Hare, "Rawls' Theory of Justice -- I" The Philosophical Quarterly Vol. 23, No.9 (1973), pp.145-146.
23. Rawls, A Theory of Justice, p.34.
24. Ibid. p.34, the footnote.
25. Hare, op.cit., p.146.
26. Rawls, op. cit., p.121.
27. Hare, op. cit., p.145.
28. Ibid. pp.145-46.
29. Peter Singer, despite the fact that he realizes this process of shuttling back and forth between our considered judgments and the principles involved in the use of the equilibrium method, still considers Rawls a subjectivist. His ground for making the claim seems to be based on the belief that only when one claims that moral principles and concepts exist objectively, can he be an objectivist. But this definition is a little restrictive. We do not really know the ontological status of logical truths and concepts, but we still firmly believe that logical truths are objective. A logician who so claims does not have to commit himself to the objective existence of logical truths and concepts. Rawls has not speculated on the problem of the ontological status of moral concepts and principles. It seems to me that Rawls' notion of the objectivity of moral principles is much more plausible when he says that the problem of the objectivity of moral principles depends on the answer to the question: "Does there exist a reasonable method for validating and invalidating given or proposed moral rules and those decisions made on the basis of them?" See Peter Singer, "Sidgwick and Reflective Equilibrium," The

Monist, 58(1974), pp.491-495 and John Rawls, "Outline
of a Decision-Procedure for Ethics," p.177.

30. Rawls, A Theory of Justice, p.149.

31. Ibid. p.587.

32. Ibid. p.579.

Part II Rawls' Ideological Framework

Chapter 5 Contractarianism and Individualism

I. An Elucidation of the Terms "Ideology" and "Individualism"

The ideological framework embodied in A Theory of Justice has been under scrutiny by philosophers of various persuasions. Liberals, radicals, and conservatives have all tried to lay bare the substantive and normative thesis Rawls wants to convey to the readers. "Justice as Fairness" has been variously labelled as "Rigid Egalitarianism," "Gladstonian Liberalism," "Liberal Individualism," and "The Contemporary Bourgeois Ideology." But in most of the literature on Rawls, he is generally considered as the champion of liberalism with an egalitarian tone. Steven Lukes' words are quite representative in this respect:

In the end, the "Archimedean point for judging the basic structure of society" that Rawls seeks eludes him. Every political theory, and every theory of justice, expresses a particular political and moral perspective. Rawls' achievement, which is considerable, is indeed to have produced a theory of justice -- a theory of liberal democratic justice. 1

In this chapter and the one follows, we shall be concerned with the issues of the ideological framework of A Theory of Justice. The approach being taken is closely related to the discussions of Rawls' moral methodology in the previous chapters. More precisely, we shall try to locate Rawls' ideological framework by way of examining the two methods (the contractarian method and the method of reflective equilibrium) that Rawls employs in constructing his theory. First, in connection with the contractarian method, we shall try to find out what is the substantive philosophical thesis that Rawls must accept before it is possible for him to use the method. It will be argued that the substantive philosophical thesis that must be accepted when the contractarian method is employed is individualism; that is to say, contractarianism presupposes individualism. The guiding idea behind this approach to the problems of Rawls' ideological framework is based on the thesis that we argued earlier, namely, that the employment of a certain philosophical method invariably commits the user to a substantive philosophical thesis. With the method of reflective equilibrium a different approach will be taken. Basically, the equilibrium method says that a set of moral principles is acceptable if and only if the principles match our considered moral beliefs in reflective equilibrium. Rawls' use

of this method together with the derivation and the acceptance of the two principles of justice in A Theory of Justice implies that the two principles match his considered moral convictions in reflective equilibrium. We shall examine the implication and significance of the acceptance of the two principles.

What is the ideological framework embodied in the two principles? What will result from the acceptance of them? And what type of society would be based on the two principles? It will be argued that the kind of society built on the foundation of the two principles of justice will be a class-divided society.

Before we launch our characterization of Rawls' ideological framework, some clarification of the meaning of the term "ideology" is needed.

"Ideology" can be defined by enumeration. It can be identified as the part of our consciousness that consists of our morality, religion, philosophy, law, etc. In other words, ideology is the part of our consciousness that Marx calls the superstructure of a culture. For Marx, the superstructure, containing morality, law, religion, etc., is in contrast with the economic substructure, i.e., the way we produce material goods. Whether the relation between the superstructure and

the substructure of a culture is essentially as Marx conceives it to be, namely, the superstructure is "determined" by the substructure and the latter always reflects the former, does not particularly concern us here.

But, from the items listed in the enumerative definition, we seem to be able to characterize ideology as:

A systematized intellectual structure of ideas which is part of the content of our consciousness. And it seems appropriate to call that part of the content of our consciousness the deep structure of our consciousness. By the deep structure of our consciousness we understand the part of our consciousness which enables us to explain the way we conceive ourselves as human beings in relation to other humans, society, and the natural world.

Ideology, when conceived in this manner, is a theoretical construct.² The holder of a certain ideology does not necessarily have to be aware of the fact that he has that particular ideology. As a matter of fact, he is not usually aware of the fact that he has a certain ideology, just as a native speaker does not necessarily have to be aware of the deep structure of the language he speaks. It suffices that when someone has successfully brought the deep structure to the surface, the bearer's overt actions and thoughts can be explained by it. Furthermore, ideology conceived in this way is primarily collective; it is ^{usually} not something which bears personal idiosyn-

cracy. Individuals obtain the ideology they subscribe to mainly through the process of socialization. The members of a collectivity share certain fundamental ideas of themselves as human beings in relation to other human beings, society, and the natural world.

Marx, being a materialist, considers ideology essentially "false consciousness" because the ideologists' approach to human reality bears an intimate relation to idealism. Both ideologists and idealists, in their survey of human reality, start from the ideas in our minds. Their approach, according to Marx, implies that ideas have an existence independent of material reality and that material reality can be understood through the grasp of ideas existing in our minds. Marx is resolutely opposed to this approach. He contends that the proper way of understanding human reality is to start from the way people actually exist, the way they work, and the manner in which they produce material goods. Thus he says:

In the first method of approach the starting-point is consciousness taken as the living individuals; in the second method, which conforms to real life, it is the real living individuals themselves, and consciousness is considered solely as their consciousness. 3

Even if the ideologists' approach to human reality is mistaken, and consequently, the ideology they articulated is

usually false consciousness, it does not entail that ideology is necessarily false consciousness. Ideology cannot be identified with false consciousness, because a materialist like Marx himself can and does articulate a proletarian ideology in the light of his observation and investigation of human reality. Of course, Marx starts his investigation from actual human existence rather than from ideas in the mind. From this, it follows that ideology is not necessarily opposed to science. An ideology does not have to be false consciousness when its carriers recognize the particularity it represents. For example, as Marx claims, the bourgeois ideology does contain certain elements of truth about the way the social system worked in the particular historical stage when the bourgeoisie was a progressive force. And this view of ideology enables Lenin to make the claim that the proletarian ideology is representative of human reality of the present historical stage. However, in most cases, ideologies do misrepresent reality and hence become false consciousness. One of the reasons for this is that ideologists usually forget that the validity of a certain ideology is confined only to a specific social class or culture. Ideology becomes false consciousness when it conceals its particularity and pretends to be universally valid.

Marx and Engels describe how this occurs in the following way.

...increasingly abstract ideas hold sway, i.e., ideas which increasingly take the form of universality. For each new class which puts itself in the place of one ruling before it, is compelled, merely in order to carry through its aim, to represent its interests as the common interest of all members of society, that is, expressed in ideal form; it has to give its ideas the form of universality, and represent them as the only rational, universally valid ones. 4

We might not agree with what Marx and Engels said in this passage about the necessity for an ideologist to claim that the ideology he subscribes to has universal validity. We also need not speculate about the intention of the ideologist in making such a claim. He might do this for a variety of reasons, for instance, self-interest, illusion, ignorance, etc. But the passage clearly points out how an ideology becomes false consciousness, i.e., how it becomes a misrepresentation of reality. An ideology becomes false consciousness when it is conceived as universally valid while in fact it is true only with regard to a certain class or culture.

In appraising Rawls' ideological framework, it seems to me that what makes Rawls' ideology false consciousness

is precisely the slip from particularity to universality. Rawls' ideology, as was stated earlier, is representative of the modern liberal democratic tradition. But he mistakenly attributes to it universal validity when he claims that an "Archimedean point for judging the basic structure of society" has been found. At the end of A Theory of Justice, Rawls says:

Once we grasp this conception, we can at any time look at the social world from the required point of view ... This standpoint is also objective and expresses our autonomy. Without conflating all persons into one but recognizing them as distinct and separate, it enables us to be impartial, even persons who are not contemporaries but who belong to many generations. Thus to see our place in society from the perspective of this position is to see it sub specie aeternitatis; it is to regard the human situation not only from all social but also from all temporal points of view.

This passage bears vivid witness to our description of how a certain ideology pretends to possess universal validity.

In this chapter we shall approach Rawls' ideological framework by way of examining his use of the contractarian method in doing moral philosophy. In characterizing his ideological framework, we shall proceed from three different perspectives. First, we shall examine what kind of ideological commitment a contractarian must make before he can make use of the contractarian device for deriving

and justifying moral principles. The claim we made earlier was that he must accept individualism. Secondly, Rawls' theory of primary goods will be examined, and it will be argued that this theory also presupposes the truth of individualism. Thirdly, contrary to Rawls' belief, we shall contend that a society based on a social contract will be individualistic. In order to substantiate this last claim, the following two points will be argued: (1) as far as their motivations are concerned, the contractors will not undergo a fundamental change after they have agreed upon certain principles to guide their conduct towards one another; (2) the people in the ordinary world will have the same kind of motivation the contractors have in the original position because the circumstances of justice will prevail both in the original position and the ordinary world. We will argue for this latter point after examining the notion of the circumstances of justice.

To begin with, we must clarify the meaning of the term "individualism." "Individualism" will be understood in two intimately related senses. And it can be shown that Rawls' theory is individualistic in both of these senses. Rawls, following Steven Lukes, defines "individualism" as

...the doctrine that the fundamental aims and interests of individuals are determined independently from particular social forms; society and the state are regarded as institutional arrangements that answer to these antecedent individual ends and purposes, as specified by a fixed and invariant human psychology. ?

Individualism, according to this definition, asserts that we can consider in abstraction an individual in the state of nature, endowed with a certain fundamental human nature, which includes his aims, needs, and interests. The aims and interests of this abstract individual are given before his socialization. Society is conceived as having only instrumental values, and its existence is solely justified on the ground that it will better facilitate human needs, aims, and interests. The abstract individual, as far as his interests, aims, and needs are concerned, will not undergo any fundamental change in joining society. Because according to individualism, all basic human needs, interests, and aims are given independently of any socialization, the doctrine will naturally deny the possibility that society can shape or reshape them. Therefore, it follows that individuals are logically and conceptually prior to society. Society contributes and adds nothing to individuals except that it satisfies the individuals' aims and interests in

a certain way. This explains why society only has instrumental value. Marx's description and criticism of individualism are so forceful that it is worthwhile quoting him at length:

The solitary and isolated hunter or fisherman, who serves Adam Smith and Ricardo as a starting point, is one of the unimaginative fantasies of eighteenth-century romances à la Robinson Crusoe, and despite the assertions of social historians, these by no means signify simply a reaction against over-refinement and reversion to a misconceived life. No more is Rousseau's Contrat social, which by means of a contract establishes a relationship and connection between subjects that are by nature independent, at all based on this kind of naturalism. This is an illusion and nothing but the aesthetic illusion of the small and big Robinsonades.... The individual in this society of free competition seems to be rid of natural ties etc. which made him an appurtenance of a particular, limited aggregation of human beings in previous historical epochs. The prophets of the eighteenth century, envisaged this individual... as an ideal whose existence belongs to the past. They saw this individual not as an historical result, but as the starting-point of history; not as something evolving in the course of history, but posited by nature, because for them this individual was in conformity with nature, in keeping with their idea of human nature. 8

The second sense of "individualism" is closely related to the first. An individualist is someone who possesses a special form of ideology. He is distinctive in his way of conceiving his relationships to other human beings, to society, and to the natural world. Essentially, in an in-

dividualist society, human beings are viewed as distinct and separate entities not having intrinsic relations with one another. The natural bodily separate existence is considered as the basis for the belief that human beings exist by themselves. In this kind of society, one is set apart from others by virtue of competing interests and goals. Individuals in society consider each other essentially as competitors for scarce goods rather than mutual helpers for realizing each other's aims. People are not dependent upon each other in any natural way because each person only cares about his own interests and goals. It is always out of consideration for their own interest that people will come to have relations with other people. Society is formed on the basis of the belief that it is to each person's interest to cooperate with others, and that rational cooperation will bring greater benefit to each and every participant. The system of market relations is the model for all other social relations: everybody engages in the activity of bargaining for maximizing his utility. Social institutions are to facilitate the smooth functioning of the market-place, and laws and morality are the enforcers of contracts. But society itself only has

instrumental value. If it were the case that everyone would be better off by not having society, then it would be disbanded. The life form in an individualist society is characterized,

...more by "having" than by "being;" i.e., your status and integrity as a person is linked to what you have (be it real property or marketable skills) rather than what you are. Property is seen as the guarantor of autonomy and well being and those without any are seen, e.g., by Locke, as deserving less than full respect and rights as human beings. Competition is accepted as inevitable based on the reality of scarcity and a belief that personal desires will always surpass what is available. 9

Moreover, human activity in the society is mostly "appropriation" and rationality is viewed as individual utility-maximizing. "Appropriating" or "possessing" rather than "appreciating" is the basic attitude people have towards the natural environment.

In the following discussion, the term "individualism" will be understood basically in these two senses.

II. The State of Nature and Individualism

Now let us proceed with the work of unpacking the ideological framework embodied in A Theory of Justice. The

first point we shall argue for is that all contractarian doctrines must presuppose individualism in both senses of the word we discussed.

(1) All contractarian doctrines must start with the notion of the state of nature. ^{9a} To consider all social relationships as contractual is to assume that we can make sense of the existence of rational beings prior to the emergence of society, inasmuch as society itself is supposed to be a state of affairs created by an agreement among the contracting parties. And since society is justified as the result of a convention among contracting parties, it is necessary that the contractors must be capable of existing prior to society. The state of nature is a pre-social state in which we consider individuals as existing separately and distinctly without any mutually agreed rules or bonds among them. Everyone is to act in accordance with the principles of his own selection. The state of nature describes the situation before any agreement is reached among the individuals. It usually consists of two sets of descriptions: One set of the descriptions tells us what the objective world is like, e.g., material goods are relatively scarce so that it is impossible to

satisfy all of everyone's desires; the other contains descriptions about the individuals themselves, e.g., their motivations, beliefs and interests, etc. The very facts that all contractarian doctrines must start from the notion of the state of nature, and that different notions of the state of nature described by various contractarian doctrines must contain a description of what human beings are like prior to the emergence of society entail that all contractarian theories must accept the proposition that individuals, with their particular characteristics, exist conceptually prior to society. The characteristics individuals have in the state of nature are determined independently of any social form. Society is defined by contractarianism as the result of an agreement among contractors.

In considering all social relationships as contractual, contractarianism is not intended to be a theory of historical explanation of how society actually emerged. Objections to this type of theory are well-known. The contractarians are not trying to tell us that various human societies all arose as the result of individuals contracting with one another. Contractarianism is concerned with the theory of justification and the rationale of social relationships

and institutions. It purports to provide a justification for the existence of society. It tries to show us why rational individuals must enter society, and why society has value for us.

From these characteristics of contractarian doctrines, it follows that the priority relation between individuals and society need not be taken by the contractarians as an historical fact. The contractarians do not have to concern themselves with the problem of how history actually developed. The alleged priority of individuals over society is a logical or conceptual requirement, i.e., in the theoretical structure of contractarianism, individuals precede society. Society is just what would be the end-result of all individuals giving their consent to a set of mutually acceptable rules for governing their conduct.

Now, the theory of the state of nature, which is the common starting-point for all contractarian theories, embodies individualism when the term is understood in the way Rawls does. Individualism is the doctrine which asserts that the fundamental aims and interests of individuals are determined independently of social institutions. The state of nature must contain a description of what these individual interests and aims are, otherwise the contrac-

10

tors would not know what to contract for. These individual interests and aims in the initial contractual situation are taken as something which are not determined by any social institutions because society comes into existence only as the result of an agreement reached by the contractors in the state of nature. In other words, the interests and aims the contractors have in the state of nature cannot be socially determined because the state of nature itself is pre-social. Where these individual interests and aims come from and how they arise are questions which the contractarians can never answer satisfactorily aside from simply asserting that they just pertain to human nature. But, of course, the opponents of contractarianism can immediately point out the difficulties involved in distinguishing between natural and institutionalized needs and interests.

The corollary that follows from the theory of the state of nature is that human beings can and must be understood apart from society. Since the fundamental aims and interests of individuals are given prior to the formation of society, it is not society which shapes or reshapes human wants and desires. Society exists simply for the

reason that it facilitates these basic human aims and interests more efficiently. It is a common belief among the contractors that everyone will be better off if all of them follow a set of mutually agreed rules in regulating their conduct towards one another. "Thus ~~man~~ is social because he is human, and not human because he is social."¹¹

Rawls is keenly aware of the fact that the theory of the state of nature presupposes the notion of an abstract individual when he says:

Now it may seem at first sight that the influence of the social system upon human wants and men's view of themselves poses a decisive objection to the contract view. One might think that this conception of justice relies upon the aims of existing individuals and regulates the social order by principles that persons guided by these aims would choose. ¹²

This passage clearly indicates Rawls' awareness of the difficulties involved in the contractarian approach to moral philosophy; yet in advocating the doctrine of primary goods, he thinks his theory can avoid the assumption that an abstract individual endowed with a particular type of interest must exist prior to society. We shall discuss this problem in the next section when we examine his theory of primary goods.

(2) To consider convention as the foundation of all social relationships and institutions entails the doctrine of abstract individuals existing prior to society. This doctrine asserts that all fundamental aims and interests of individuals are given independently of social institutions. The contractarians think that individuals can and must be understood apart from their social existence and that human sociability is the product rather than the condition of social existence. All this is tantamount to saying that by nature human beings are not sociable, and that sociability is not part of the intrinsic nature of human beings, since if human beings were naturally sociable, it would not be necessary for them to contract with one another to form a society. To be non-social would be contrary to their nature. It follows that individual consent, which is considered the binding force of all social relationships, would become redundant, because it would be natural for human beings to come together and form societies. But, the contractarians consider social contract and individual consent as necessary ingredients for the emergence of society. These elements are artificial in the sense that they do not necessarily come into existence. The only t

thing that pertains to human beings naturally, according to contractarianism, is their never-ending quest for the satisfaction of their own interests. The existence of society is justified solely on the ground that it will promote everyone's interests better. We can look at the problem from another angle. If it were easier to satisfy and execute one's interests and aims in the state of nature than in society, it would follow that individuals would find no reason to contract with one another to form a society, because one enters society simply out of the belief that society can facilitate the satisfaction of one's interests more efficiently. Individual aims and interests existed prior to society; they are not derived from or created by society. The correlate of this thesis that human beings are not naturally social is, of course, the doctrine that society only has instrumental value. It is not an intrinsic good in itself. The reason for its existence is merely that it enables us to satisfy our interests to a greater extent. Contractarianism justifies society solely on the ground that its existence is to everyone's advantage.

(3) Since according to contractarianism, society is not an end in itself, it follows that its existence can

only be justified on the ground that it embodies instrumental value: i.e., wanting to satisfy our wants and desires we find that society, with a set of mutually accepted rules of conduct, will enhance our interests better than if we all remain in the state of nature. To view the function of society in this manner will impose some restrictions on the way the contracting game is played, and it will also put some severe restrictions on the sort of goods which the contractors in the state of nature bargain for. Since social cooperation is considered mutually advantageous, the contracting game cannot be a strict zero-sum one. In a zero-sum game, one wins only when others lose. If it were a zero-sum game, the persons who lose would not consider it worthwhile to play the game. The kind of game the contractors are willing to play must be what game theorists call negotiable games. In a negotiable game, each player believes and expects that the outcome of the game is beneficial to everyone, even though each player is concerned only with his own interests and gains. The restrictions the instrumental notion of society puts on the sort of goods people in the state of nature seek are rather severe. They can not be the sort of goods that are strictly competitive, i.e., the goods cannot be of such

nature that no increase of supply is possible as the result of social cooperation. The sort of goods, being strictly competitive, would entail that the sort of game they were engaged in would be a zero-sum game, and mutual improvement becomes impossible. The result will again lead those who lose to refuse to play the game.¹³ For instance, there are two individuals, A and B, and two hundred apples. If A and B are to act according to the principles of their own selection in attempting to obtain the apples, A will get 150 apples and B only 50, because A is smarter and stronger than B. Moreover, suppose the supply of apples will not be increased through the cooperation between A and B. Now B proposes to play a zero-sum game, and according to the rules of the game, the outcome is that B gets 65 apples and A's share is reduced to 135 apples. In this case, A will certainly prefer the state of nature in which he gets 150 apples and refuse to play the game with B.¹⁴ So, the restriction imposed on the sort of goods can be sought in the state of nature by the contractors is that the goods must not be strictly competitive. Secondly, the sort of goods cannot be completely non-competitive, either; i.e., the sort of goods cannot be of such nature that the

increase of the goods for some people will not in any way affect other's share. If this were the case, the contractors will find no reason to cooperate with one another. Also the goods cannot be of unlimited supply in the state of nature and in society. If this were the case, again, no mutual improvement is possible through cooperation. Thus, the contractors will find it unnecessary for them to play the game. The condition of relative scarcity of the goods must prevail both in the state of nature and in society. All these conditions impose certain restrictions on the sort of goods which contractors in the state of nature bargain for. The goods that satisfy these conditions will be goods: (1) whose supply can be increased for everyone through cooperation so that each will be better off than he is if he remains in the state of nature; (2) whose supply cannot be so abundant so that the problem of distributive share will not arise. In spite of the possibility that mutual improvement can be achieved through cooperation, it is still true that frequently the increase of the goods for one person will affect others' share of the pie.

What sort of goods can satisfy these two conditions? Social goods such as friendship, love, etc. can be ignored at this moment. Intellectual goods such as knowledge, though

their amount can be increased through cooperation, do not seem to be such that their possession by an individual A excludes possession by another individual B. So, it does not satisfy the second condition above. In other words, it will not necessarily create a problem of distributive justice. It is true that contractors are willing to cooperate in obtaining knowledge, since its supply can be increased by cooperation, but since the share of knowledge does not create a problem of distributive justice, the contractors will have no reason to include knowledge among the list of the goods they would bargain for in the state of nature.

Rawlsian primary goods, with probably one exception,¹⁵ will satisfy the two conditions. The list of primary goods includes rights and liberties, opportunities and powers, income and wealth, etc. The supply of these goods can be expanded through cooperation, and yet their quantity is not so abundant that it does not create a problem of distributive shares.

The restrictions imposed upon the sort of goods the contractors will bargain for in the state of nature lead us to the conclusion that each person's good is distinct

from and even opposed to the good of others. The kind of goods which can be bargained for by the contractors in the initial contractual situation are goods the appropriation of which by one person always excludes the appropriation of them by others. For instance, if A has appropriated a certain amount of food, it will be impossible for another individual B to possess the same food without A's losing it. In a situation of relative scarcity, which prevails in the state of nature, each person treats everyone else as competitors for these precious goods. This is the state of hostility as described by Hobbes, Locke, and Rawls. And the state of hostility as a description of the state of nature is the common starting-point shared by various forms of contractarianism. Rawls is right when he says that it would be useless to ascribe perfect altruism to the contractors in the initial contractual situation. If everyone is perfectly altruistic, the result will be the same as if everyone is self-interested, since an altruist insists on what he proposes to be done, i.e., to do whatever others want.¹⁶ Secondly, in using a contractarian approach for doing moral philosophy, one of the pre-requisites is that the contractors are not previously tied by any moral feel-

ings toward each other. The ascription of altruism to contractors violates this pre-requisite.

The nature of the goods the contractors can bargain for in the initial contractual situation makes it necessary for contractarians to presuppose the truth of individualism. An individualist is characterized by the way he conceives his relationships to other human beings, to society, and to the natural world. As has been stated earlier, in an individualist culture, people are viewed as separate and distinct beings. Their interests are also distinct and often opposed. It will never occur to an individualist that his existence is possible because it is interwoven with other people's existence. An individualist will always adopt what Derek Parfit calls the "simple view" of personal identity.¹⁷ His personal identity is taken to be as deep as possible. And in an individualist society, people's lives are generally characterized more by "having" than by "being" or "appreciating." This is the kind of ideology all contractarian theories must presuppose in the beginning. Due to the peculiar nature of the goods the contractors can bargain for in the state of nature, all of them are appropriators of these goods for themselves. "Appropriating" rather than "appreciating" is the basic

form of existence for the contractors. And since everyone's goods are opposed to the goods of others, each must view himself as existing distinctly and separately. No bonds or natural ties exist between them whatever aside from the relationships established by contract. Rawls accuses the classical utilitarianism of not taking seriously the distinctions of persons. He thinks this failure one of the main drawbacks of utilitarian doctrine. Justice as Fairness avoids this drawback by assuming individualism. A utilitarian can also charge Rawls with not taking seriously the fact that our existences are mutually interwoven.

We have seen why all contractarian doctrines must presuppose individualism. Being a contractarian doctrine, Justice as Fairness is no exception.

III. The Theory of Primary Goods

In this section, we shall examine Rawls' theory of primary goods. It will be argued that the theory does have individualistic connotations in the sense that by using it to derive the principles of justice, Rawls has made some commitment to a notion of an abstract individual endowed with certain fundamental aims and interests. The claim to be contended for here is that the theory of primary

goods presupposes individualism in the sense Rawls understands the word. It should not be confused with the claim that the theory of primary goods entails a set of principles of justice which in turn leads to an individualistic society. The theory of justice derived from the theory of primary goods together with other premises may or may not lead to an individualistic society, but it is a totally different issue than the one that concerns us here.

Rawls claims in "Fairness to Goodness" that the account of primary goods is not individualistic.¹⁸ But it seems that he confuses the two different problems we distinguished above. He claims that if the principles of justice derived from the theory of primary goods and other premises do not lead to an individualistic society, then the theory of primary goods itself is not individualistic. Thus he states:

Effective arguments for the conclusion that the account of primary goods is individualistic would have to show at least the following: first, that in a well-ordered society effectively regulated by the principles of justice people would acquire and pursue individualistic ways of life.... Second, one must show that the reason for these results is the account of primary goods. The argument has to be that if this account were changed and different principles were adopted, the corresponding well-ordered society would not be individualistic. 19

Then he proceeds to argue that a well-ordered society based on his two principles of justice will not be individualistic because it stresses that the interests and aims of individuals in that society depend on existing institutions and the principles of justice they satisfy.²⁰ But, aside from what Rawls pointed out in the passage quoted above, there is another way of showing that the theory of primary goods is individualistic. If we can show that the theory requires a prior commitment to a certain notion of an abstract individual endowed, independently of social institutions, with certain fundamental aims and preferences, then it follows that individualism must indeed be presupposed. Further, if we can show that the theory of primary goods is not neutral in the sense that it is not universally applicable to human beings of different cultures, and the items included in the list of primary goods are not equally useful to different kinds of life project, then we will be able to refute a very important claim made by Rawls, namely, that the premises he has employed in the original position are weak and widely accepted. Also, it will follow that the claim that the device of the original position is fair becomes dubious.

What is Rawls' theory of primary goods? To put it simply, it asserts: first, there is a set of things which a rational person wants whatever else he wants, in order to carry out his life plan, whatever it may be. The set contains things such as rights and liberties, powers and opportunities, and income and wealth, etc. Second, rational contractors in the original position would prefer more rather than less of these goods.

From the very beginning, Rawls rejects teleological theories as unacceptable. This outright rejection is based on the contention that teleological theories take existing human desires and aims as given. In a teleological theory, the concept of right is defined in terms of the concept of non-moral good; the concept of non-moral good is defined independently, either naturalistically or non-naturalistically. The device of the original position effectively blocks contractors' knowledge of their specific aims and interests. Now a list of primary goods is needed in the original position as the index of expectations for the contractors who are situated behind the veil of ignorance and deprived of the knowledge of their desires and aims; otherwise they would have no way of knowing what they should bargain for.

Although the contractors do not know their specific desires and interests, they nevertheless know that they will have certain life projects when they enter the ordinary world. If we deprive them of all their ideas of what goodness is, they might find that the principles they consented to in the original position are not only unhelpful to their life plans, but also destructive. This is one of the reasons why a theory of primary goods is required in the original position. The theory enables the choosing game to get started. Secondly, if the contractors in the original position are completely ignorant of what things are good for them or what things they want, they will have no way of knowing what the distributive principles they try to reach are about. The distributive principles must concern the distribution of something. Since they cannot be about the distribution of utility conceived as the satisfaction of existing desires and wants inasmuch as Rawls rejects teleological theories, the primary goods, which are supposed to be wanted by all rational people, become the natural candidates. This is another reason why the theory of primary goods is introduced.

In Part One of A Theory of Justice, when Rawls tries to derive the two principles of justice from the original

position, he simply asserts that: (1) there are primary goods and they are the things rational people would want whatever else they want in order to carry out their life plans; (2) the contractors in the original position would prefer more rather than less of the primary goods; (3) primary goods include rights, liberties, income, wealth, etc. He did not argue for any of the claims then. However, Rawls realizes that these claims need to be substantiated, and a theory of the good is required to justify them. Indeed, a theory of the good is provided for in Part Three of A Theory of Justice to account for the theory of primary goods. Rawls calls this explaining theory "the thin theory of the good," and contrasts it with "the thick theory of the good." The theory is called the "thin theory of the good" because it does not involve the concept of right in any way. In contrast, the "thick theory" is supposed to explain various sorts of moral goodness by using the thin theory together with the two principles of justice already derived from the original position. Basically, the thin theory does not assert specifically what people's aims and desires are; instead, it only says that people's wants and desires have a certain structure. The distinction between the thin and the thick theories of the good is cru-

cial to Rawls. The theory of primary goods must be explained by some sort of theory of the good, and the explaining theory itself cannot involve any controversial moral ideas, otherwise the theory of justice derived from it will be based on some morally non-neutral concepts. As a result, the theory of primary goods will also lose its moral neutrality. In other words, the list of primary goods will not be universally valid for all people if the thin theory of the good is not morally neutral and condones certain moral conceptions and rejects others. For this reason, only a thin theory of the good which does not contain any controversial elements as to whether certain desires or wants are morally favored or repugnant can serve the purpose.

The thin theory of the good can be taken as equivalent to the three-stage definition of the good given by Rawls: (1) An object is good if it has the properties which are generally desired in objects of its kind; (2) an object is good for a person if it has the properties he rationally desires in objects of its kind; (3) an object is good for a person if it has the properties he rationally desires in objects of its kind and the person's desire is part of a system of desires which belongs to his plan of life, and

the plan of life itself is rational. As to what is a rational plan of life, Rawls has the following to say.

... 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 full deliberative rationality, that is with full awareness of the relevant facts and after a careful consideration of the consequences. 21

The principles of rational choice are, according to Rawls:

(1) The principle of taking most effective means to given ends.

(2) The principle of inclusiveness -- one plan is preferred to another if its execution can achieve all the desired aims of the other plans and one or more further aims.

(3) The principle of greater likelihood -- one plan is preferred to another if the likelihood of its success is greater than other plans.

For our purpose, we do not need to attack these principles directly. By the same token, we need not consider directly whether the three claims Rawls makes about the primary goods can be successfully explained by the thin theory of the good. Nevertheless, if we can show that the

theory of primary goods, which is derived from the thin theory of the good, is problematic, then we will have every reason to doubt the plausibility of the thin theory itself.

Of course, our main purpose is to show that the theory of primary goods is individualistic, i.e., it makes a commitment to a certain notion of an abstract individual endowed with certain fundamental aims and preferences. We will be successful if we are able to show that the theory is not universally applicable to people with different cultural backgrounds. One way to show this is to reveal that the motivational structure of some human beings does not exhibit the characteristics described by the theory of primary goods. This can be proved if we can show that the claim that it is natural for all contractors to want more rather than an adequate amount of the primary goods is false, and that to have more primary goods is not really conducive to some types of life projects. Consequently, it follows that after all it is not so natural for all human beings to want more rather than less of the primary goods; rather, it is accidental and peculiar to certain type of society and culture. Our argument against Rawls' claims about

primary goods parallels C.B. Macpherson's argument against Hobbes' claim about human nature in the state of nature, i.e., that man is to seek ever more power. Macpherson points out that it is not natural for man to seek ever more power; it is merely accidentally true of man of a certain type of society. The argument we shall offer against Rawls' claims about the primary goods is that there are some types of life plans which will be counter-examples to the claim that rational persons will always desire more rather than a certain adequate amount of the primary goods.

To illustrate our case, first, we shall borrow an example from Adina Schwartz.²² Let us imagine a rational person somewhat in the line of the early Marx. The young socialist has certain beliefs, for example, that political structures are "determined" by economic structures; that only a non-market economy can achieve equality for all; that a good life must rest on self-realization through meaningful works, and so forth. He also firmly believes that the life plan that he has based on these convictions does not presuppose the preference for more rather than less of the primary goods, because he is convinced that to center his life plan on endless appropriative activities

is a great hindrance to the realization of having a life of meaningful work. As a matter of fact, to prefer and possess more rather than an adequate amount of the Rawlsian primary goods will be for him not only not conducive to his life plan, but also damaging to it.

Let us also imagine a rational individual whose life plan is that of a Buddhist monk's. He believes in the basic Buddhist teachings of the Four Noble Truths: (1) that all life is sorrowful; (2) that sorrow is due to craving; (3) that sorrow can be overcome only by stopping craving; (4) that this can be achieved only by following a life of carefully disciplined conduct, which culminates in the life of meditation. He further believes that to appropriate more material goods than the bare essentials necessary for his physical existence is not only not conducive to his life plan, but also harmful to it. Human relationships based on a contract, to him, are not genuine. And our relationship to the natural world as a conqueror to the conquered is basically wrong.

Both the young socialist and the Buddhist monk will agree with Rawls on the claim that they need a certain amount of the primary goods in order to sustain their lives and

to carry out their life plans. But they will dispute with Rawls about the claim that people will necessarily want more rather than less of these goods in the original position. They will argue that there are certain types of life plans whose successful execution does not presuppose that to have more primary goods is always better than to have less of them. They will both claim that it is irrational to want more rather than less of the primary goods as long as we can have an adequate amount of them. To possess more material goods is not really useful to their life plans. On the contrary, it will distract them from the kind of life projects they intend to pursue. Rawls' reply that they can discard the extra primary goods they do not need as long as they want to after the veil of ignorance is lifted is not really to the point here. What they intend to point out is that the theory of primary goods presupposes that to possess more rather than less of the primary goods is always conducive to every type of life plan, and the presupposition is unwarranted. Both the Buddhist monk and the young socialist only want to point out that there are certain types of life plans that are incompatible with this assumption. Rawls' appeal to Marx, by arguing that even Marx would want the primary goods, and that the only thing

distinctive about the socialist's pursuit of the primary goods is that he wants them to be controlled collectively is off the mark here. Our young socialist is not disputing with Rawls about the claim that the Rawlsian primary goods are useful things for carrying out his life plan. The point he tries to make all along is simply that it is not necessarily irrational to want less than more of the primary goods. Conversely, it is not necessarily rational to want more rather than less of the primary goods.²³

To presuppose ~~that~~ to have more rather than less of the primary goods is always good for every type of life plan is certainly dubious, as our two examples have shown. However, Rawls' theory of primary goods does make this dubious assumption. Rawls is certainly right in claiming that what motivations people will have in the ordinary world depends on what type of society they live in.²⁴ But our point is simply to make it clear that by using the two examples, it can be shown that the theory of primary goods is not necessarily true of everyone in every society. To derive a theory of distributive justice on the account of primary goods is to base it on very shaky ground. The thin theory of the good might not be so thin after all. And

since the theory of primary goods is not compatible with certain types of life plan, it does condone only certain types of life plans which consist of certain particular aims and desires. The incompatibility between the theory of primary goods and certain types of life plan entails that the theory is not impartial with respect to different types of life plan. The partiality of the theory of primary goods further entails that it assumes that human motivation has a certain structure. Any type of motivational structure that is not compatible with what the theory of primary goods specifies will be considered irrational. From these, it follows that the theory of primary goods does make a prior commitment to a notion of an abstract, supposedly universal human nature, which is in fact not common to every human being. And, given how the theory has to specify this human nature, it can be concluded both that the theory of primary goods has individualistic connotations, and that it is not universally applicable.^{24a}

IV. The Circumstances of Justice and Individualism

...radical contractarianism is incompatible with the view that men undergo fundamental change in becoming members of society. Men's reasons for

contracting one with another are supposed to arise out of their presocial needs in the state of nature. If contractarian ideology is to be effective in rationalizing social relationships, then these needs must be represented, not only as presocial, but as permanent, so that the reasons for entering the contract will also be reasons for maintaining the society created thereby.

David Gauthier

This section can be taken as an elaboration on the statements made by Gauthier.

In the previous two sections, we have argued that contractarianism must presuppose the truth of individualism. First, we discussed generally the claim that all contractarian doctrines must make the commitment to the truth of individualism because they all must start with the notion of the state of nature, and this notion has individualistic implications inasmuch as it contains a notion of an abstract individual. Then we discussed Rawls' theory of primary goods, and argued that the theory also must assume individualism since it is incompatible with certain types of life plans. This fact implies that a notion of an abstract individual which is peculiar to certain type of society and culture is embodied in the theory of primary goods. But we did not argue then that contractarianism will also lead

to individualism, i.e., that basing all social relationships on contracts will lead to individualistic societies. In this section, we shall argue that contractarianism, which considers all human relationships as conventional, will indeed have this implication. A society based on a social contract will be a civil or private society in which people's main concern is the promotion of their own interests. We shall argue for this point by examining Rawls' notion of the circumstances of justice. Essentially, the structure of the argument will be as follows: The circumstances of justice obtain only if people are mutually disinterested; when all human relationships are conceived as based on a social contract, the circumstances of justice will inevitably obtain. Furthermore, the circumstances of justice will prevail not only in the original contractual situation, but also in the ordinary world. And it is a sufficient condition of a society's being an individualistic society that it is one in which people take only their own interests into consideration. Therefore, contractarianism inevitably leads to individualism.

Rawls essentially follows Hume in his account of the notion of the circumstances of justice. The problem of

distributive justice arises, according to Rawls, only when the circumstances of justice obtain. He divides the conditions which constitute the circumstances of justice into two parts. First, there are the objective conditions which make human cooperation both possible and mutually advantageous: many individuals coexist together at the same time in a definite territory. No one is powerful enough either physically or intellectually to force other people to do whatever is advantageous to him. Natural resources are not so abundant that cooperation becomes superfluous. Secondly, on the subjective side, people are mutually disinterested, i.e., they take no interest in one another's interests. Yet, they have roughly similarly needs and interests, or needs and interests in certain ways complementary to one another. They are rational in the sense that they are individual utility-maximizers. Rawls puts special emphases on the conditions of moderate scarcity among the objective conditions, and mutual disinterestedness among the subjective ones.

From the brief descriptions of the conditions constituting the circumstances of justice, it seems to be the case that the two conditions of moderate scarcity and mu-

tual disinterestedness are jointly sufficient for the problem of distributive justice to arise. Whenever we have both of them present, i.e., when the natural and other resources are not so abundant that it is possible to satisfy all of everyone's wants, and people only care about the satisfaction of their own desires and take no interest in the satisfaction of other people's desires and wants, then we will face a problem of distributive justice. Also, each of these two conditions is necessary for the problem of distributive justice to occur. So, even if the condition of moderate scarcity obtains, if people have strong altruistic feelings towards each other and generally take other's interests into consideration, then the problem of justice will not arise. By the same token, even if the condition of mutual disinterestedness obtains, the problem of justice will not occur if we have infinite resources so that everyone's desires and wants can be satisfied. The obtaining of the circumstances of justice implies that both the conditions of moderate scarcity and mutual disinterest prevail. Hume puts the point very succinctly as follows:

For what purpose make a partition of goods where everyone has already more than enough?... why call this object mine when the seizing of it by

another, I need but stretch my hand to possess myself of what is equally valuable? Justice, in that case, being totally useless, would be an idle ceremonial and could never possibly have place in the catalogue of virtues.

Again: suppose that, though the necessities of the human race continue the same as at present, yet the mind is so enlarged and so replete with friendship and generosity that every man has the uttermost tenderness for every man, and feels no more concern for his own interest than for that of his fellows: It seems evident that the use of Justice would, in this case, be suspended by such an extensive benevolence, nor would the divisions and barriers of property and obligation have ever been thought of. 25

Hume's arguments here are essentially similar to the points we made above. Each of the conditions of moderate scarcity and mutual disinterestedness is a necessary condition of the circumstances of justice in which alone the problem of distributive justice will arise. The obtaining of the circumstances of justice implies that both of the conditions are present. 25a

In order to establish our thesis that a well-ordered society based on Rawls' two principles of justice will be an individualistic society in which people are motivated by self-interest, we need to show that the circumstances of justice will obtain in that society. If the circumstances of justice obtain in a Rawlsian well-ordered society, it will follow that people in that society are mutually disinterested.

In dealing with the problem of people's motivation in Rawls' theory, we must first make a distinction between the motivation the contractors have in the original position and the motivation people will have in a Rawlsian well-ordered society. As to the problem of who are self-interested, the contractors in the original position or the people in the ordinary world, Hare points out that Rawls is ambiguous on this issue.²⁶ When Rawls ascribes the motivation of self-interestedness to people, he sometimes can be taken to refer to the contractors in the original position, and sometimes to the people in the ordinary world. He warns us explicitly not to confuse the latter's motivation with the former's. However, it is quite clear from certain passages in A Theory of Justice that the contractors in the original position are self-interested:

For the fact that in the original position the parties are characterized as not interested in one another's concerns does not entail that persons in ordinary life who hold the principles that would be agreed to are similarly disinterested in one another. 27

But the postulate of mutual disinterest in the original position is made to insure that the principles of justice do not depend on strong assumption. 28

But whether people are either mutually disinterested or otherwise in ordinary world after the veil of ignorance

is lifted is not so explicitly explained by Rawls. Some passages in A Theory of Justice seem to suggest that people in everyday life are as mutually disinterested as they were when they were situated behind the veil of ignorance.

But a human society is characterized by the circumstances of justice. 29

...I have assumed all along that the parties know that they are subject to the conditions of human life. Being in the circumstances of justice, they are situated in the world with other men who likewise face limitations of moderate scarcity and competing claims. 30

But, on the other hand, from Rawls' discussion of the idea of a social union in Section 79 in A Theory of Justice and the article "Fairness to Goodness," we seem to be assured that Rawls thinks that people in a well-ordered society based on his two principles will not be individualists, i.e., they will not be primarily self-interested as they were in the original position. He says:

There is no reason why a well-ordered society should encourage primarily individualistic values if this means ways of life that lead individuals to pursue their own way and to have no concern for the interests of others (although respecting their rights and liberties). 31

Our contention will be that there is good reason to suppose that a well-ordered society based on a social contract will be a private or civil society. And the people

in that kind of society who conceive all social relationships as contractual, will be individualists who will primarily pursue their own ways of life and have no concerns for the interests of others even though they might respect each other's rights and liberties.

A theory of social justice is, to a great extent, based on a certain specific notion of society. Different conceptions of society will result in different conceptions of justice. Opposing theories of justice will view society in an opposed manner. They will differ in their views about the functions of society, the human relationships people have in society, the foundations of society. For example, Rawls' theory of justice basically takes society to be a voluntary cooperative venture for mutual advantage, and it is typically marked by a conflict as well as a partial identity of interests of its members. Classical utilitarianism will view society mainly as a collectivity; it is more or less like a family in which personal interests among members are interwoven, and personal identities are not taken to be so distinct and deep. Rawls accepts this view of the correlation between a conception of society and a theory of justice.

The various conceptions of justice are the outgrowth of different notions of society against the background of opposing views of the natural necessities and opportunities of human life. Fully to understand a conception of justice we must make explicit the conception of social cooperation from which it derives. 32

Principles of justice are, according to Rawls, identified by the role they play in society. They are the principles which will determine the various distributions of advantages of social cooperation; they will assign duties and rights in the basic institutions of society; and they will also define the appropriate distributions of benefits and burdens for each member who participates in the social cooperation. Rawls attaches enormous importance to the virtue of justice. He actually considers justice the first virtue of social institutions. No matter how efficient or stable a society may be, and how well-coordinated it is, if it is unjust, it will be in need of urgent reform.

Now, the great need for the principles of justice in a society to adjudicate conflicting claims is a definite proof that the circumstances of justice obtain in the society. To place as much importance on the virtue of justice as Rawls does clearly implies that the principles of justice play a correspondingly important role in that society. These claims can also be corroborated by the view on the foundation of society Justice as Fairness takes.

Society is a cooperative venture for mutual advantage, however, even if it is mutually advantageous for the members of a society to cooperate in the joint venture, a society based on a contract is still marked on many occasions by conflict of interests among its members. The principles of justice are needed not only in designing the basic structure of a society; they will continually be appealed to when conflicts actually occur in society. They will be used to adjudicate conflicting claims when these occur in everyday life. The very fact that they will continually be required in society is a clear evidence that the circumstances of justice obtain in ordinary world. If it were not the case that the circumstances of justice obtain, there would be no need for the principles of justice. And the fact that the circumstances of justice obtain in a Rawlsian well-ordered society is enough to show that in that society, people are mutually disinterested. It has been established earlier that mutual disinterestedness is a necessary condition for the obtaining of the circumstances of justice. So, we may conclude that in a Rawlsian well-ordered society, where the virtue of justice occupies such a centrally important role, people generally tend to be individualists who will care very little about the interests of others.

Rawls tries to argue that even when the circumstances of justice obtain, people are not necessarily self-interested. He gives as an example the relationships between members of different religions. They are supposed to be mutually disinterested under the circumstances of justice, yet they are neither self-interested nor necessarily committed to the pursuit of individualistic plan of life. Two remarks can be made about this example. First, when we say that the persons belonging to religious group A are not self-interested, we mean that they work together for the same goal within the group. Within the boundary of the group itself, the members are not pursuing their own individual goals; instead, they have a common end which they share with each other. Of course, the common end is identical with the end each participant strives for. Under this condition, the problem of justice will not arise since conflict of interests will not occur. To be sure, there will be conflicts, but the nature of conflict is different from the sort of conflicts which occur in an individualistic society in which people are pursuing their own goals. The sort of conflict which might occur in a religious group will be conflicts of opinions as to the best way of achiev-

ing their common goal. The notion of distributive justice does not really apply here. But the relations between members of different religious groups will be a completely different story. If they are under the circumstances of justice, there will be conflicts of interests. They will need principles of justice to settle their conflicting claims, e.g., which group should have more access to the public forum. This fact entails that they are self-interested when they come to have contact with each other on occasions of relative scarcity. Unlike the relations between the members within a religious group who share a common goal, the members of different religious groups do not have a shared goal which all of them pursue.

Locke's theory of the two types of social relationships is of some help in explaining the differences between these two types of relationships between people of the same religious group and people of different religious groups. He thinks that the relations among members of the same family are not contractual; they are conceived as in accordance with the natural law. The problem of justice does not arise there. But when the head of the household leaves the house and does business with someone else in society, the relation-

ship will be contractual. The circumstance of justice will prevail there; and a set of principles of justice based on a social contract will be needed in regulating their conduct towards each other.

These two remarks show that Rawls' example is not really convincing as far as his answer to the problem of whether people will be self-interested in the circumstances of justice is concerned. Self-interestedness, being a necessary condition for the circumstances of justice to obtain, must prevail whenever the latter obtain. The very fact that Rawls puts so much emphasis on the virtue of justice strongly suggests that he conceives society to be a place where conflict of interests occur frequently, and the motivation of self-interestedness occupies an eminent place in the society. Of course, this is not to say that self-interestedness can disappear totally in some other type of society, whose foundation is other than a social contract, self-interestedness might occupy only a subordinate place or even be reduced to an insignificant amount.

To argue that society is based on a social contract is essentially to view people living in the society as having, among other things, opposing interests. The reason they

are willing to enter a society is that it will make each and everyone of them better off than if they were to stay in the state of nature. They believe that their personal needs can be better advanced by a cooperative venture. The important thing is that they must believe that their personal needs and desires in the state of nature will not change upon entering society. If those personal needs were to undergo fundamental changes after entering society, the contractors will have every reason to suspect whether the arrangement of social institutions is conducive to their newly acquired needs and desires which they did not have in the state of nature. They would have no reason to prefer one kind of social arrangement to another since they do not know anything about what they will want in the society. And since they do not know what desires they will come to have, on what basis they can choose one social arrangement and reject another? To argue that human motivation will undergo drastic change upon entering society is to downgrade or even nullify the significance of making a social contract. Rawlsian contractors, situated behind the veil of ignorance, are deprived completely of the knowledge of their particular interests and desires. They do not know

what things to bargain for due to this ignorance. The theory of primary goods is introduced to rescue this impossible situation. The primary goods are supposed to be universally wanted whatever life plans the contractors will have in ordinary life. By making this presupposition, i.e., that the theory of primary goods is universally applicable to different types of life plans, Rawls has made the primary goods the thread which connects the desires and interests of the contractors in the original position and the desires and interests they will come to have upon entering the actual world. But he has also tacitly conceded the point we have argued for, namely, that the presocial needs and desires the contractors have in the original contractual situation will not undergo change upon entering society.

Footnotes

1. Steven Lukes, "An Archimedean Point," Observer Review, 4 June, 1972.
2. Here, we can compare our conception of ideology with Rawls' conception of moral philosophy. Actually, my conception of ideology as the deep structure of the content of our consciousness is inspired by Rawls' comparison between moral philosophy and linguistics. See A Theory of Justice, (Cambridge, Massachusetts: Harvard University Press, 1971), pp. 46-47. However, the things we have to note here is that Rawls thinks that the deep structure of our moral capacity does not differ between different people in different societies and cultures. Our conception of ideology does not commit us to this position.
3. K. Marx and F. Engels, The German Ideology, Part One, with selections from Part Two and Three and Supplementary Texts, ed. with introduction by C.J. Arthur, (New York: International Publishers, 1970), p.47.
4. Ibid. p.547.
5. John Rawls, op. cit., p.587.
6. Rawls argues in "Fairness to Goodness" (Philosophical Review, Vol. 84 (1975), pp.536-554) that the theory of primary goods is not individualistic. See especially pp. 546-47.
7. Ibid. p.547.
8. K. Marx, Introduction to a Critique of Political Economy, in German Ideology and Supplementary Texts, ed. by C.J. Arthur, p.124.
9. Larry Elle, "Individualism and Class Consciousness," Radical Philosopher's Newsjournal, (IX) Fall 1977, p.2.

9a. The reason that all contractarian doctrines must start from the notion of the state of nature is that by using a social contract as the justification of the existence of society, the contractarians must view society as something which is the result of an agreement among the contracting parties. Since society is considered as the result of people coming together for an agreement, within the theoretical structure of contractarianism, there must be a distinction between the pre-social and the social. Society emerges only after an agreement is reached. The state of nature is just the pre-social stage of life. However, contractarianism does not require that actual human beings temporarily precede society, what it requires is simply a logical or conceptual priority of the individuals over society. All three versions of contractarianism, namely, overt, tacit, and hypothetical, must start with the notion of the state of nature. Even though hypothetical contractarianism does not make any commitment to the existence of an actual contract, and it is only a sort of "thought experiment," it is still necessary that in the "thought experiment" the state of nature precedes society. Consequently, the pre-social state wherein human beings live without a set of rules commonly agreed upon is prior to society. In other words, human beings must be posited as existing prior to society.

It might be argued that what we said above is not true of Rawls' theory since his task is to select a set of principles of justice rather than to establish a society. However, we can certainly view Rawls' theory as a sort of social theory. Any normative theory of society can be viewed as an attempt to provide a set of principles on the basis of which society should and can be built. The two principles of justice chosen in the original position will be the organizing principles of a Rawlsian well-ordered society. In view of Rawls' stress on the subject matter of his theory, namely, the basic structure of society, we are certainly justified in making the claim that his hypothetical contractarianism is an attempt to provide a foundation for the establishment of a society.

10. This is one of the reasons why Rawls has to provide a theory of the good in A Theory of Justice. The result is the theory of primary goods which we shall examine later.
11. David Gauthier, "Social Contract as Ideology," Philosophy and Public Affairs, Vol. 6, No. 2 (Winter 1977), p.138.
12. Rawls, A Theory of Justice, p.260.
13. For the discussion of the restrictions on the sort of goods the contractors in the state of nature can bargain for, I am heavily indebted to David Gauthier. See his article "Social Contract as Ideology," op. cit., especially pp. 144-48. He tries to show there that if we conceive all social relationships as contractual, we will be led to the thesis that human activities are mainly appropriative.
14. Of course we have to assume that A is a utility-maximizer that he prefers more apples to fewer.
15. The item I have in mind is self-respect. It seems that self-respect can be increased in two ways: (1) one's own sense of the worth of one's own life plan; (2) other people's respect for the value of one's life plan. (1) is not something that can be bargained for in a contracting game. Even though through other people's recognition, one's self-respect can be increased, it seems to me to be more appropriate to say that the contractors will include among the items of primary goods "mutual respect" rather than self-respect. Also, self-respect will not typically create a distributive problem. The possession of self-respect by one person does not exclude the possession of the same amount of self-respect by others. Later, in "Fairness to Goodness," Rawls talks about the social bases of self-respect rather than self-respect as a primary good. This seems to corroborate my view that what the contractors seek is mutual respect rather than self-respect. I am not trying to say that self-respect is unimportant. On the contrary, I believe it to be the most important thing for a fulfilling life. What I am saying is simply that self-respect is not the sort of thing that will by itself create any distributive problem.

16. For Rawls' discussion of perfect altruism, see A Theory of Justice, p.189. The point he makes there is that unless someone has independent, first-order desires, no one can get his desire fulfilled in a society of perfect altruists.
17. Derek Parfit, "Later Selves and Moral Principles," Philosophy and Personal Relations, ed. by Alan Montefiore, (Montreal: McGill - Queen's University Press, 1973), pp. 137 - 169.
18. Rawls, "Fairness to Goodness," see especially pp. 544-46.
19. Ibid. p.544.
20. Rawls' arguments for this claim are not at all convincing. Despite the fact that he discusses the idea of a social union at length (see A Theory of Justice, Section 79), one still very much wonders whether the state of affairs described there can be derived from his theory of justice. As a matter of fact, we shall argue later that since the circumstances of justice prevail both in the original position and in the ordinary world, a Rawlsian well-ordered society based on a social contract will essentially be an individualistic or civil society. Also, see his discussion about how society based on his two principles will be divided into non-comparing groups. (A Theory of Justice, p.545).
21. Rawls, A Theory of Justice, p.408.
22. Adina Schwartz, "Moral Neutrality and Primary Goods," Ethics, Vol.83, No.4 (July 1973), pp.302 - 304.
23. Rawls, "Fairness to Goodness," p.541.
24. Schwartz's example of the young socialist is illuminating, but it does not serve her purpose. She claims that the young socialist can be justified in claiming that he will be hurt in a society based on a preference for more rather than less of the primary goods:

Our socialist could object to Rawls' argument by claiming that he would be harmed by living

in a society based on a preference for a greater rather than lesser amount of wealth. He could say that, living in such a society, he would devote valuable time to thinking about material wealth and trying to decide whether or not to avoid the temptation of attempting to acquire more possessions. (Adina Schwartz, op. cit., p.304.)

She only presupposes that people in a Rawlsian well-ordered society will prefer more rather than less primary goods, but this claim needs to be substantiated. We shall try to substantiate this point in the next section when we examine the notion of the circumstances of justice in Rawls' theory.

- 24a. The point we try to make in this section about the individualistic implications of the theory of primary goods can further be strengthened by pointing out that: First, the theory of primary goods is used by Rawls as premises in the contractarian argument for the derivation of the principles of justice, therefore, it describes human beings in a pre-social state. The motivational structure of human beings described by the theory of primary goods cannot be the result of socialization inasmuch as society has not been established. Second, the theory of primary goods can be based only on the thin theory of the good otherwise the former will lose its neutrality with respect to various moral conceptions. The primary work of the thin theory itself is to describe the motivational structure of human beings who have not entered society yet. In fact, society has yet to be established by the principles of justice which are to be derived from the original position. In *Justice as Fairness*, the thin theory is logically prior to the principles of justice. The theory of primary goods, which also logically precedes the principles of justice, is therefore employed to describe the motivational structure of the contractors in the initial contractual situation. Whether the thin theory of the good and the theory of primary goods can adequately describe the motivational structure of human beings in society depends on what kind of society it is.

25. David Hume, An Inquiry Concerning the Principles of Morals, ed. with an Introduction by Charles W. Hendel, (New York: The Bobbs-Merrill Company Inc., 1975), pp.15-16.
- 25a. The contention that the scarcity of material goods is only a necessary but not a sufficient condition for the problem of justice to arise can be shown in the following manner: it is quite possible for a society to experience material scarcity without at the same time facing the problem of justice provided that human beings have extensive benevolence and are willing to give the scarce material goods to the most needy ones. In that society, people might encounter problems of how to determine who are the most needy ones, but this is not a problem of distributive justice. Of course, in that situation, there will also be problems of how to distribute goods, but not every problem of distribution concerns justice. Rawls agrees with us on this point that the scarcity of goods does not necessarily create problems of justice. He gives an example of a society of saints, presumably it is not rich, in which everyone works selflessly for the glory of God. In that society, there will still be conflicts, but the conflicts will be conflicts of opinions as to how to serve God best rather than conflicts of interests. See A Theory of Justice, pp. 129-130. It has been argued that an ideal Marxian society is also a society which is beyond justice. For this point, see William L McBride, "The Concept of Justice in Marx, Engels and Others," Ethics, Vol. 85, No.3 (April, 1975), pp. 204-218; Michael Teitelman, "On the Theory of the Practice of the Theory of Justice," Journal of Chinese Philosophy, 5(1978), pp.217-247.
26. R.M. Hare, "Rawls' Theory of Justice," in Reading Rawls, ed. by Norman Daniels (New York: Basic Books Inc., 1975), p.87, the footnote.
27. Rawls, A Theory of Justice, p.148.
28. Ibid. p.129.
29. Ibid. p.136.
30. Ibid. p.257.
31. Rawls, "Fairness to Goodness," p.550.
32. Rawls, A Theory of Justice, pp.9-10.

Chapter 6 The Difference Principle and the Differences of Life-Prospect

I. The Derivation of the Difference Principle

A work on Rawls is not complete without an examination of the principles of justice themselves. These principles are the substantive matters of Rawls' theory of justice as fairness. According to Rawls, they are the principles we should use for setting up a well-ordered society. The two principles are:

1. 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.
2. Social and economic inequalities are to be arranged so that they are both:
 - (a) to the greatest benefit of the least advantaged, consistent with the just savings principle, and
 - (b) attached to offices and positions open to all under conditions of fair equality of opportunity.

According to the method of reflective equilibrium, a moral theory is justified if and only if its principles match our considered moral convictions in reflective equilibrium. Since this method is one of the two methods Rawls employs in constructing and justifying his theory, it must be the case that the two principles he derives from the

original position match his considered judgments in reflective equilibrium. It is Rawls' conviction that a society based on the two principles will be a just society.

In this chapter, we shall examine Rawls' substantive normative theses. In particular, we shall concentrate on the difference principle. Two specific questions will be raised and examined in detail: (1) Is the derivation of the first part of the second principle, i.e., the difference principle, from the original position valid? (2) Is this principle really an egalitarian principle as Rawls claims it to be? It will be our contention that the difference principle is not logically derivable from the original position, and that the contractors in the original position, given the assumptions about their motivation, will not necessarily select it over some other principle. In order to make the derivation more plausible, an additional premise is needed. We shall argue further that even with this additional premise in hand, Rawls still will not be able to make the derivation of the difference principle from the original position logically necessary. Indeed, we shall argue that it is highly probable that the contractors will reject the difference principle. With regard to the

claim that Justice as Fairness is an egalitarian theory, we shall argue that due to a fundamental assumption made by Rawls, i.e., that inequality is inevitable in any society, his claim cannot be substantiated. His application of the idea of pure procedural justice to the problems of distributive justice is problematic and the kind of society based on his two principles will be a class-divided society in which the total life prospects of the children of entrepreneurs and children of unskilled laborers will be considerably different, even when these children are endowed with equal talents and abilities. This result will greatly decrease the plausibility of Rawls' theory since it is his firm conviction that in a just society, the distribution of goods should not be dependent on the fortune of natural endowment and social contingency.

Characteristically, all liberal political theories try to provide a two-fold justification. On the one hand, liberal theorists try to argue for various political liberties for all citizens; on the other hand, they attempt to provide a justification for various sorts of economic inequalities. The two-fold justification of equal political liberties for all citizens and economic inequalities among them is based on a fundamental assumption that equal

political liberties and economic inequalities are compatible. Different liberal theorists advocate different sets of equal basic liberties as well as different sorts and degrees of economic inequality. For example, Locke argued for a set of equal liberties of political participation, and Mill tried to defend equal liberties of thought and expression. Usually, they base their justification of equal political liberties on a doctrine of natural rights or social contract, or even on utilitarian calculation. But, when it comes to the problems of the distribution of material goods, liberal theorists believe that economic inequalities of wealth, power, authority, and income can be justified because people are endowed with different talents, merits, etc., and they also make different amounts of contributions and efforts to the production of material goods.

The central theme of A Theory of Justice is concerned with this traditional liberal two-fold justification. The first principle deals with the problem of how various basic political liberties should be distributed. As can be expected from any liberal theory, Justice as Fairness advocates an equal distribution of various basic rights among all citizens. The difference principle (the first part of the second principle) is concerned with the problem of how exten-

sive economic inequalities in income, wealth, power, and authority can be allowed and justified. The second part of the second principle (the principle of fair equality of opportunity), being a modification of the difference principle, stipulates that fair equality of opportunity should be maintained in the process of the distribution of various material goods.

In examining the plausibility of a theory of justice and the principles it advocates, the most important things to see are: (1) from what assumptions are the principles derived; (2) how are they derived; (3) what kind of society will emerge when it is based on the principles in question. The principles viewed in isolation will only give us part of the picture. As we know, Rawls' principles of justice are derived from the original position which consists of a set of statements describing the original contractual situation. Rawls considers the relation between the original position and the two principles of justice a logical deduction. In examining the normative theses of Justice as Fairness, we shall concentrate our discussion only on the derivation of the difference principle. As was said earlier, our contention will be that in order to make its

derivation more plausible, Rawls must add at least an extra premise about human motivation. The additional premise required is that contractors in the original position are materially motivated. Material incentives are the major forces behind their acceptance of the difference principle which allows and even requires economic inequalities among citizens.

Rawls' arguments for the acceptance of the difference principle rely on three basic assumptions: (1) contractors in the original position are mutually disinterested; (2) they want as many primary goods as possible; and (3) they are devoid of envy. The primary goods are rights and liberties, opportunities and powers, and income and wealth. From these three basic assumptions, Rawls argues that the difference principle will be selected:

But an equal division of all primary goods is irrational in view of the possibility of bettering everyone's circumstances by accepting certain inequalities.²

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.³

First of all, we have to make clear what Rawls means by "inequalities." It seems reasonable to assume that in

modern industrial societies, in order to feed, clothe, and accomodate the multitude of population properly, some sort of division of labor is required so that production can be carried out more efficiently to meet people's needs. Different people in society must play different roles to achieve this goal. It is also plainly true that various social roles, in order to carry out their functions properly, will have to carry with them different amounts of authority and power. In modern industrial societies, we need engineers, lawyers, bus drivers, farmers, policemen, and politicians, etc. Different kinds and amounts of power and authority attached to these various social positions are necessary for the efficient exercise of the jobs. For example, a policeman must have a certain kind and amount of power and authority in order to enforce the laws effectively; a senator must have a different kind and amount of authority and power in order to make effective legislation which concerns the welfare of all people. As long as a division of labor remains indispensable, we will require different people to play different social roles; and as long as there are different people playing different social roles, the differences in authority and powers necessary to effectively carrying out the various jobs will remain. We are not here to make any judg-

ment as to whether or not, in the future, division of labor will be totally eliminated or reduced to an insignificant degree. But it seems reasonable that we should not consider as inequalities the sort of differences in authority and power created by the necessity of division of labor. The reasons for this is that on the one hand, division of labor is required at least at this stage of human history for the maintenance of a certain standard of living for the multitude of the population and on the other hand, the differences of authority and power attached to various social positions are considered necessary for the effective functioning of the positions. They are not artificially attached privileges which have no internal relation with the positions. If these differences of power and authority are taken as inequalities, it is hard to imagine how human beings, given that we need a certain degree of industrialization, and with industrialization comes division of labor, can ever achieve equality.

In his essay, "Justice as Fairness," Rawls makes it clear that he does not consider the kind of differences in authority and power created by the various social roles as inequalities. Inequalities occur only when we attach external-

ly different rewards to different social positions. Inequalities are not mere differences of offices and position, rather, they are the by-product of holding different offices and are in a certain sense, artificially attached to the offices and positions. For instance, in order to carry out effectively the functions of the office of the Prime Minister of Canada, it can be argued that, given today's technology and standard of living and the complexity of the job, the occupant of the office, should have an airplane. So, in a sense, the privilege of having an airplane at his or her disposal, in the capacity of being the Prime Minister of Canada, is a necessity which goes with the office of the Prime Minister. The difference between the Prime Minister and other citizens, as far as having an airplane is concerned, is a difference necessitated by the nature of the office. The difference here should not be viewed as inequality.

But it seems to be arbitrary to say that in order to do the Prime Minister's job efficiently, his salary should be ten times that of a professor's. If, as a matter of fact, the Prime Minister's salary is actually ten times higher than a professor's, the difference between them in this respect will be considered as an inequality. Since the reward to the office of the Prime Minister is not necessary for the

effective exercise of the job, it is not a necessary and integral part of the office and position itself. It is rather artificially and arbitrarily attached to the office.

The difference principle tries to justify inequalities of this kind which are not mere differences. Now let us consider why in deriving the difference principle from the original position, the additional premise -- that the contractors are materially motivated -- is required. When rational contractors have settled on the first principle of equal liberty,⁴ "as well as an equal distribution of income and wealth," why would they go a step further and deviate from this state of equality? It seems that the assumptions of self-interest, the pursuit of the primary goods, together with the lack of envy, are not sufficient for them to do that. Material incentives on the part of the contractors must also be assumed. Even if the contractors are mutually disinterested and try to gain as many primary goods as possible for themselves and are not vulnerable to envy, it does not logically follow that they will deviate from the state of equality and accept an unequal distribution of wealth and income if they do not consider material goods more valuable than the things which the state of equality can bring about.

e.g. a strong sense that one's own worth is not lower than anyone else's. By adopting the first principle, the contractors have obtained equality not only in liberties, but also in other areas, such as power and authority, etc. It is true that "equality" is not one of the items included in the list of the primary goods. But it is still something the contractors consider to be valuable. In arguing for the difference principle, Rawls states that "the immediate gain which a greater equality might allow can be regarded as intelligently invested in view of its future return."⁵ The implication of this statement is that equality is a state of affairs which is of utility to the contractors. The "future return" referred to here clearly means the increased material goods that can be obtained by allowing inequalities of goods. It seems more plausible to think that the contractors behind the veil of ignorance will not deviate from the initial state of equality achieved by adopting the first principle if they do not consider that the gain on the material goods, which presumably can be obtained by inequalities, outweighs the good things equality can bring about.

Rawls argues that when it comes to the problem of the distribution of liberty, the contractors will not allow any-

one to have more basic liberties than others, even if by doing so all will have more liberty; furthermore, the contractors will not trade the equality of basic liberties for more material gains, i.e., they will not accept less than equal liberty in order to have more material goods, (this is what the priority of the first principle over the difference principle means). But why would they trade the equality of material wealth for more material wealth in absolute terms and take the risk that they might end up having less of the material goods, than someone else? The assumption, that the contractors have a desire for having ever more primary goods in itself, cannot explain it. The basic liberties are also primary goods. Why are the contractors not willing to trade the equality of basic liberties for more liberties in absolute terms and taking the risk of having less liberties than someone else? The very fact that they will insist on equal basic liberties for everyone, instead of unequal liberties among them with the provision that everyone's basic liberties be increased because of the inequalities, shows that the contractors consider equality of liberties more important than the increase of liberty in absolute terms. This fact also shows the inadequacy of the attempt to use

the assumption about the contractors' desire for having more primary goods to explain their willingness to accept inequalities in income and wealth, etc. In order to make the derivation of the difference principle more plausible, the additional premise, that the contractors are at least partially materially motivated when they join the cooperative adventures for selecting a set of principles to govern their conduct towards each other, is required. This additional premise, that the contractors are materially motivated, will explain why the contractors are willing to give up equality of material wealth for more material goods in absolute terms.

The difference principle has been severely criticized by many writers.⁶ Its plausibility depends on Rawls' contention that it will be irrational on the part of the contractors to want an equal distribution of material goods instead of bettering their positions on wealth in absolute terms. But is it really irrational to prefer equality of the distribution of material wealth to unequal distribution together with a possibility of improving one's own situation in material wealth? Rawls' notion of instrumental rationality at least does not direct us clearly either way. It seems

that in order to understand clearly how Rawls justifies the acceptance of the difference principle, we have to investigate his ideas of "better off" and "improving one's position" a little deeper. He understands them in purely monetary terms. In arguing for the difference principle, Rawls maintains that the contractors will prefer more material goods to equality of goods for all, as long as the inequality will increase one's share. But as has been argued earlier in connection with our criticism of the theory of primary goods, it does not seem plausible to claim that rational people will always want more rather than a certain adequate amount of the primary goods which will guarantee that their plans of life can more or less be successfully carried out. The purpose of the assumption that people will always want more rather than an adequate amount of primary goods is, according to Rawls, to set "... up a determinate choice problem by supposing that the parties maximize something (in particular primary goods)."⁷ But, as Rawls himself stresses, the conditions imposed on the original position must be reasonable, true, and widely accepted; one cannot simply include certain premises in the description of the original position for the sake of expediency or to make one's account come out

in the way one wants it to. This is exactly what happens here when, due to the necessity of solving a problem of determinate choice, Rawls introduces some implausible premises into the description of the initial contractual situation. In particular, the assumption that the contractors will always prefer more rather than an adequate amount of primary goods. As a consequence of this implausible assumption that the contractors will always prefer more rather than an adequate amount of primary goods, Rawls' understanding of the terms "better off" and "improving one's situation" becomes problematic. To understand the notions of "better off" and "improving one's situation" in the Rawlsian manner does not always seem to square without considered judgments about values. They are understood rather too narrowly and distortedly. Kai Nielsen puts the point as follows:

The rub, however, is in Rawls's understanding of 'better off' or 'improving the position' of the worst off. He cashes these notions in purely monetary terms. This prompts the response that either this is too narrow a notion of being 'better off' or of 'improving your position', or we are not justified in believing that rational agents, who have a tolerably adequate conception of fairness, will always give first priority to being 'better off' or 'improving their position'. They might very well, in conditions of moderate scarcity, recognize other things to be of greater value.

Furthermore, in view of the possibility that inequality of material wealth tends to create inequalities of other things, such as liberties, rights, opportunities, and self respect, etc., the contractors will be very cautious with the move of allowing inequality of the distribution of material wealth.⁹

"Envy" is another thing we have to take into account when we consider the narrowness of Rawls' ideas of "better off" and "improving one's position." Rawls contends that the insistence on strict equality may derive from envy.

To be sure, there may be forms of equality that do spring from envy. Strict egalitarianism, the doctrine which insists upon an equal distribution of all primary goods, conceivably derives from this propensity.¹⁰

If this is really the case, i.e., if envy underlies demands for equal distribution of all primary goods, then since envy itself is a destructive force which tends to make everyone worse off, we have no choice but to abandon strict egalitarianism. But as Lawrence Crocker has argued convincingly, envy is not the sole possible basis of strict egalitarianism.¹¹ It is quite conceivable that people possess a set of dispositions which he calls "solidarity"; those endowed with this kind of disposition tend to have strong

feelings of mutual identification, cooperation and fraternity. They all view themselves as part of the team where all of them participate in some joint adventure. The inability to tolerate the difference of wealth among the members and the different life prospects created by the inequalities does not arise so much because of envy. It is rather because they are afraid that the sense of mutual identification and fraternity might be lost because of the differences in wealth and life prospects among various members. It is also highly plausible, that a person endowed with a strong sense of justice will not tolerate the disparities of life prospects caused by inequalities of wealth and income among the people of a community. This person does not necessarily have to be in the position of the least advantaged class. He might even be a member of the ruling class, as Engels was. It is his strong sense of justice which makes him believe that it is unjust for people to have unequal life prospects, and it is this strong sense of justice which makes him advocate egalitarianism. It seems then, that strict egalitarianism does not necessarily derive from the destructive propensity of envy. It can be based on such healthy feelings as "solidarity" and "justice." The Buddhist maxim that "if

there is one person in the world who is not liberated, then I will not be liberated" expresses this sort of feeling of solidarity. The narrowness of Rawls' conception of "better off" prevents him from seeing that there might be something which is of higher value for which the contractors will strive. It is not really irrational to give up some material goods in order to obtain equality among all human beings.

II. The Maximin Rule and the Derivation of the Egalitarian Principle

It has been argued in the previous section that:

- (1) in order to make the derivation of the difference principle more plausible, Rawls needs at least an extra premise, namely, that the contractors are materially motivated; and
- (2) Rawls' notions of "better off" and "improving one's situation" are rather narrowly conceived. His notion of "rationality" is, to say the least, contestable. To choose equal shares rather than more material goods in absolute terms, for the contractors, is not necessarily irrational. Now we shall push a step further and contend that even if the extra premise about the contractors' motivation is included in the original position, the difference principle still will

not necessarily be chosen by the parties to the contract. The argument for this contention will be as follows: we shall agree with Rawls, for argument's sake, that the maximin rule is the most reasonable criterion to adopt for making choices under uncertainty. We shall then argue that by following the maximin rule, it is at least as rational for the contractors to prefer the egalitarian principle of equal shares to the difference principle as the other way around.

When making choices under uncertainty in which we have no way of calculating the probabilities of the possible outcomes, the maximin rule instructs us that "we are to adopt the alternative the worst outcome of which is superior to the worst outcomes of the others."¹² The rule becomes the best strategy for decision-making when three features are present in a situation of choice:

"First, since the rule takes no account of the likelihoods 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." ¹³

Rawls believes that the original position embodies all three features. Whether or not it does, is a matter

of dispute, but that does not concern us here.

It has been argued earlier that since, by assumption, the contractors in the original position are motivated only by self-interest, and that due to the device of the veil of ignorance, they are prevented from knowing all particular facts, e.g. who they are, what their interests and desires are, and what talents they have, etc., consequently, they totally lose their self-identities. But it is unreasonable to ask the self-interested egoists, i.e., the contractors, to adopt the general point of view when making choices of the principles, because, by definition, a self-interested individual is a person who will not take up the universal or general point of view. Unless there is some way out of this predicament, the contracting game cannot be carried out. However, one way out of this plight is by providing the contractors a way which enables them to identify themselves temporarily in the original position with a certain social group in order that the choice of principles can be made. The major purpose of the adoption of the maximin rule is precisely to achieve this. Rawls argues that by not allowing the contractors behind the veil of ignorance to have the knowledge of who they are, they will naturally adopt

a conservative approach to decision-making, and therefore, for the purpose of making choices, temporarily identify themselves with the least advantaged group of the social strata. The contractors figure that since it is not likely that they can know the possible circumstances and the probable outcomes of their choices, they had better be careful and play it safe. Elaborating in this manner, they think that it is in their interest to maximize the interests and utilities of the worst-off group in society. It is only by adopting the maximin rule that the contractors are able to find their indentities behind the veil of ignorance and consequently make their choices. Again, it is a matter of controversy as to whether or not the maximin rule is the best strategy for decision-making under uncertainty, and we shall leave it aside.

The point we intend to argue is that even supposing that the maximin rule is the best strategy for decision-making under uncertainty, contrary to Rawls' belief, the contractors will not necessarily prefer the difference principle to the principle of equal shares. The derivation of the difference principle from the original position is not perfectly sound.

When self-interested people gather together to make a contract, they believe that they will become better off

by cooperating with one another, and they figure that through cooperation, their situations will be improved (here, let us understand "better off" and "improving one's position" in the manner Rawls does, i.e., in purely monetary terms). The subject matter with which the difference principle deals is the issue of how to distribute justly material goods among the contracting parties. In considering the problem of social cooperation where the question of distributive justice of material goods arises, there are at least three possible cases that have to be taken into account, as far as the supply of the goods, which is the result of cooperation, is concerned.

(1) The amount of material goods is more or less fixed, and these goods are already present and they are not the result of anyone's work. In this case, nobody can claim any credit for the production of the goods. An example of this situation is if a group of people jointly discover a bushel of apples in the forest. Nobody has actually produced them since they are wild apples. There is no problem of who makes more contributions or efforts for the production of the goods.

(2) The second possible case is that through cooperation, the total amount of material goods would be less than

the amount they would have produced if the contractors were to act alone and without cooperation.

(3) The third possible case is that as a result of cooperation, the amount of material goods will be increased. In this case, the contractors will face the problem of how to distribute adequately the goods generated by their joint efforts, unlike case one, in which the elements of each and every contractor's contributions and efforts are involved in the production of the goods they now must distribute among themselves. And these factors will render the distributive problem much more difficult.

Case two can be dealt with easily, since a decrease of the total amount of goods will mean that some people will have to settle for less than they would have been able to obtain if there had been no cooperation. And given that contractors are self-interested and that they are also utility-maximizers, cooperation will not be in the interest of those who end up having less than if they had acted alone. Consequently, they have no reason to make a contract with other people.

Case one is also relatively easy. If the total amount of material goods is fixed and cannot be increased through cooperation, then, given that the contractors are situated

behind the veil of ignorance, that they do not know who they are, it seems plausible to assume that they will settle for equal shares. This seems to be what the maximin rule will prescribe, and indeed this is the solution Rawls advocates:

Now actually, if ... the stock of goods were generally thought to be given, then a strict opposition of interests would be assumed to obtain. In this case, it would be correct to think that justice requires equal shares. Social wealth is not viewed as the outcome of mutually advantageous cooperation and so there is no fair basis for an unequal division of advantages. 14

Case three is the situation in which all contractarian doctrines find they have a serious problem. It involves the belief that society is a mutually advantageous venture in which the problem of distributive justice arises. The situation described in case three is a partial description of the original position where the Rawlsian contractors find themselves. They also believe that cooperation is in their interest. We have to add some other conditions to make it resemble more the description of the original position. The contractors are situated behind the veil of ignorance, hence they are deprived of all knowledge about particular facts. Now Rawls argues that under this situation, the contractors will adopt the maximin rule as their guide for

making choices of the principles of justice, and that the maximin rule will dictate that they settle upon the difference principle. The argument for this claim, as we have seen above, is that it is irrational to settle for equality in view of the possibility of improving one's own situation if inequality is allowed. Our contention is that even if the contractors adopt the maximin rule, they will not necessarily choose the difference principle.

Let us suppose there are ten people. When there is no cooperation among them, the total amount of goods produced by them separately amounts to \$1,000. Let us call this sum S_1 . They come together with the belief that if they cooperate, the total sum of material goods produced jointly will be increased. And let this sum be S_2 . Suppose its worth is equivalent to \$3,000. \$2,000 worth of goods are made possible due to cooperation among the contractors. Now they face the problem of how to distribute "the goods produced by them jointly." There are two possible ways they can deal with this problem: (1) they can agree that the principles of distribution apply to the total sum S_2 jointly created by them; or (2) they can agree that the principles of distribution apply only to the incremental amount $S_2 - S_1$, i.e., the surplus which was the result of their cooperation. ¹⁵

Rawls does not distinguish these two possible ways of dealing with the problem of distribution. He is only concerned with the first case, i.e., how the total sum S_2 is to be distributed. The case he discusses seems to be more realistic. It is rather unlikely that a person would know the amount of goods that he could produce if there had been no cooperation, since no one had ever been in the state of nature. Also, someone might argue that the total amount S_2 which is jointly produced by all the participants is much greater than the total sum S_1 which is not reflected in the figures of S_1 and S_2 in our example. Therefore, S_1 is rather insignificant. Now suppose in the state of nature, the worst-off individual can produce by himself an amount of wealth equivalent to \$80 (where the total sum by adding up each person's production is S_1 , i.e., \$1,000.), and the better-off individual can make \$120. With the intention of improving their situations (in purely monetary terms), they decide to cooperate. There are a number of principles available to them as a guide to how they might distribute the goods produced jointly. Let us just consider two of them: (1) the difference principle; (2) the egalitarian principle of equal shares.

According to the difference principle, if the worst-off individual, by cooperating with other people, can obtain \$150 worth of material wealth while a more advantaged person gets \$400 worth of goods, the former should accept the inequality allowed by the principle (in this case, the total sum is S_2 , i.e., \$3,000.); since the inequality is required to extract the efforts of the more advantaged person in order that the total amount of goods can be increased and so too the worst-off person's share will be increased to \$150. It would be irrational for the worst-off individual to refuse the inequality since if he does, he would end up getting less than \$150. This is the whole rationale behind the acceptance of the difference principle. But it seems to be defective.

The least advantaged individual believes that: (1) the maximin rule is the correct criterion to apply in the original position; (2) everyone is self-interested; (3) the contractors are materially motivated; and (4) by cooperating with each other, the total amount of goods will be increased, although he does not know the exact amount of increment which will result from the cooperation. Given these beliefs, we can see the least advantaged individual reasoning in a dif-

ferent logic from the one which leads to the acceptance of the difference principle. This way of reasoning is at least as sound as that leading to the difference principle.

	X	Y
A	100	100
B	150	400
C	300	300
D	200	200
E	120	120

In the table above, X, Y are individuals, X is the least advantaged person in terms of ability and talent required for securing material wealth, and Y is a more advantaged person. A, B, C, D, and E are possible patterns of distributing different amounts of wealth among ten people (the total amount of wealth in A is \$1,000., in B and C, it is \$3,000., in D, it is \$2,000., and in E, it is \$1,200.). We only have to consider the situations of two individuals, namely X and Y. In A, the total amount of wealth (\$1,000) is the sum of goods produced by ten individuals separately. When they settled upon the first principle of equal shares of everything, the resulting distribution is that everyone got \$100. Now, due to cooperation, it is possible to increase the total amount of wealth to \$3,000. The contractors are faced with

the problem of choosing a principle of distribution to share this new wealth. Rawls argues that they will choose the difference principle which can be used to justify B in which the least advantaged person, X, gets \$150, while Y gets \$400. Our contention is that the contractors will prefer the egalitarian principle of equal shares to the difference principle, X will figure that if he accepts the difference principles, then it is possible that he would end up getting \$150 while some other person would get \$400 (as in B). But if he insists on the egalitarian principle, that everybody shares the total sum equally, then he would get \$300 worth of material wealth (as in C). Of course, he will take into consideration, the fact that people are materially motivated; if the more advantaged individual is not allowed to have \$400, he might not put in as much effort as he would have if inequality is permitted, hence the total amount produced jointly would not be as high as the total sum of B, (i.e., \$3,000 which is the amount they will produce jointly when the difference principle is adopted as their guide for distributing the goods). In that case, the amount he would get would not be as high as \$300 (which is the result of following the egalitarian principle of dividing the total sum of B among ten people). It might be only \$200 (as in

D), or even less than \$150 (as in E) which is the amount he would get if they all had accepted the difference principle. But he knows that the more advantaged person is also materially motivated. Even if the egalitarian principle is adopted, for the sake of his own benefit, the more advantaged person would still want to produce as much as possible, since by so doing, his share would be increased. Even though his reward for putting in more effort in this case might not be as high as in the case of following the difference principle.

By following the egalitarian principle, the probability that the total sum they will produce jointly will exceed \$1500 (in which case, the worst-off will get more than \$150 which is the amount he will get by adopting the difference principle; in fact, by adopting the principle of equal shares, everyone will get the same amount) is at least as high as the probability that the total amount will be less than \$1500 (in which case, the worst-off will get less than \$150). So the choice between the difference principle and the strict egalitarian principle is not really settled. It is by no means clear that the difference principle will be preferred. Given the facts that the more advantaged person is also ma-

terially motivated, and that he does not suffer from jealousy and envy, it seems very likely that he will put in as much effort as he can so long as the effort he puts in does not reach the point where the gain offsets the effort. Moreover, since the contractors situated behind the veil of ignorance are all to adopt the maximin rule and to identify themselves with the least advantaged group of society, then they will likely try to maximize the utility of the worst-off individual as much as possible. Given that the probability calculation of the possible outcomes is rather vague, choosing the egalitarian principle is at least as good as adopting the difference principle. Further, by assuming that the contractors are not vulnerable to jealousy and envy, it becomes more probable that the total amount that will be produced by following the egalitarian principle will be higher than \$1500, since it is possible to produce \$3000 by their joint efforts. The more advantaged individual will not be jealous simply because the least advantaged individual is getting more, as long as the gain which the more advantaged individual receives does not offset the effort he puts in. In the final analysis, the problem of the principle to be chosen by the contractors, whether the difference

principle or the egalitarian principle, depends on the point where the gain offsets the effort people put in. But this is an empirical question affected by many factors; it varies from society to society and time to time and cannot be settled in advance.

Admittedly both the difference principle and the egalitarian principle will prefer C to B in the table above, but the point X will keep in mind is that it is possible to use the difference principle to justify B, in which he will get less than someone else. The egalitarian principle can never be used to justify B. If the argument that it is very likely that the total amount of wealth produced by joint effort will exceed \$1500 is sound, it is safer for X to choose the egalitarian principle over the difference principle. In addition, when the contractors take into consideration the result that the egalitarian principle will bring about, e.g., economic equality in addition to political equality, equal liberty and more importantly, equal worth of liberty, the discontinuation of divisions into non-comparing classes, and most importantly, equal self-respect,¹⁶ then the case of the selection for the egalitarian principle is greatly enhanced.

III. The Difference Principle and Class-Division

Liberty and equality have been the declared social ideals of all liberal theories. Political theorists in this tradition are dedicated to the task of justifying these two social ideals. By liberty, they mean various political rights; and equality, for them, specifies equality of opportunity and equal political rights for all citizens. The justification of these two social ideals is also the central concern of A Theory of Justice. Thus, Rawls says,

...we can associate the traditional ideas of liberty, equality, and fraternity with the democratic interpretation of the two principles of justice as follows: liberty corresponds to the first principle, equality to the idea of equality in the first principle together with equality of fair opportunity, and fraternity to the difference principle. 17

However, when we actually encounter the classics of the liberal tradition, it is a little puzzling that one of the fundamental concerns of all liberal theories has been the justification of inequalities of the distribution of economic wealth among citizens. In undertaking this task, there are two fundamental beliefs all liberals hold. First, they believe that economic inequalities are inevitable in any kind of society and that disparities of life prospects

among citizens, created by economic inequalities are compatible with what justice requires. Second, they are convinced that equal political liberty and economic inequalities are not incompatible. According to liberals, a person with less means to the fulfillment of his or her human capacity can enjoy the same amount of political rights as a person who possesses more means to this fulfillment. Even liberal theorists with different ethical backgrounds, e.g., a utilitarian liberal such as J. S. Mill, or a contractarian liberal such as Locke, all share these convictions. Rawls, being a liberal, is no exception.

In spite of his claim in the quotation above, that equality is a social ideal obtained in Justice as Fairness, it still remains that one of the central concerns of A Theory of Justice is the justification of unequal life prospects for members of different social strata. According to the difference principle, economic inequalities are to be allowed, or even required, if they are to work to the advantage of the least favored social class. Economic inequalities sanctioned by the difference principle, Rawls claims, are not only justified, but also just. The fundamental reason for making this claim is that this type of

economic inequality is presumably unavoidable in any society. Rawls says,

... consider the distribution of income among social classes ... those starting out as a member of the entrepreneurial class in property-owning democracy, say, have a better prospect than those who begin in the class of the unskilled laborers. It seems likely that this will be true even when the social injustices which now exist are removed. What, then, can possibly justify this kind of initial inequality of life prospect?¹⁸

Rawls believes that which justifies the inequalities in life prospects between different social classes is the difference principle. But, as C. B. Macpherson points out,

"this principle makes sense if social and economic inequalities as between classes, amounting to inequality in life prospects, are presumed to be 'inevitable' or even 'likely' to persist even when existing social injustices have been removed." ¹⁹

He then goes on to challenge Rawls on the soundness of the fundamental assumption that inequality and class-division are inevitable:

Is a classless society unthinkable? Is it, that is to say, impossible to envisage a society in which, even if there are perceptibly different levels of income and authority, the occupancy of a higher level is neither the result nor the means of exploiting others (in the strict sense of exploitation, i.e., transferring to oneself for one's own benefit some of the powers of others)? I do not think so. But I see a possible reason why Rawls is unable to envisage it, namely that he does not see that class division in any society, not least in his free market

society, is based on such continuous transfer; the transfer is the means and the result of class division. 20

Clearly, from this passage, we can see that Macpherson's notion of class division is considerably different from Rawls' notion of class division. Macpherson understands class division in the traditional Marxist sense, i.e., that class division is a result of the institution of private property and exploitation (transferring to oneself for one's own benefit some of the powers of others); while Rawls understands class division in terms of contemporary sociological theory of social stratification. Rawls may respond to Macpherson by saying that in the Marxist sense of class division and inequality, it is indeed not impossible to achieve classlessness.²¹ But the question remains: is class division and inequality in the Rawlsian sense, i.e., institutionalized inequality which amounts to different life prospects for different social classes inevitable? As far as we can ascertain by examining the texts, Rawls has not given any argument to support or to make unproblematic the assumption that class division (in his sense) is inevitable. The upshot of this assumption is that we just have to accept the differences in life prospects between the children of entrepreneurs and unskilled laborers. The inequalities referred to here are

neither mere differences of offices and positions, nor differences of talents and abilities. Nobody expects division of labor to disappear in the near future, nor will anyone deny that it is likely that people will continue to have different natural assets even in a classless society. The inequalities are the disparities of life prospects between different social classes. It is a result of the way in which we organize our society; some can develop their human capacity more fully than others because of the access or the lack of access to the necessary means for this development. The disparities do not have to be there, they are purely man-made. Differences in natural endowments do not necessarily create inequalities, nor do differences in social roles and positions. Only when we decide to use these differences as bases to attach artificially different rewards, do inequalities emerge. In order to justify the assumption that classlessness is impossible, Rawls at least must offer some arguments to the effect that it is impossible to make a social system work when all the disparities in life prospects are removed. Since he has not done this, the plausibility of the difference principle, which heavily depends on this assumption, becomes more problematic.

Marx considers the institution of private property and division of labor as the two major causes of class division and exploitation. Private ownership of the means of production, as Macpherson would say, is the major cause and means of "transferring to oneself for one's own benefit some of the powers of others." The owners of the means of production, simply by owning the capital and employing laborers, are able to collect the surplus value created by the workers. But, clearly, it is not necessary that there be private ownership of the means of production. Rawls concedes that his theory of justice is compatible with socialism where the means of production is publicly owned.²² It seems that the main reason Rawls believes that there will always be class division is that he thinks that division of labor is inevitable. As was stated earlier, in an industrial society, in order to feed, clothe, and accommodate the huge population, the efficiency made possible by division of labor seems necessary. But it seems that division of labor is only a necessary condition of class differentiation or inequality. Division of labor tends to create class division only when it is practiced in a certain type of society. Traditionally, division of labor goes hand in hand with the institutionalized

23
social statuses. There are certain social roles which always carry with them high prestige, income and wealth, etc., and other social positions are usually connected with low prestige, income and wealth, etc. But, as has been argued before, the connections between various social positions and high income, wealth and prestige are artificial in the sense that the positions do not necessarily go with the goods. The so-called good things are not an integral part of the positions, rather, they are the result of holding these positions when they are rewarded or regarded in a certain way by a certain social system. The connection between them is purely man-made. The high prestige, income and wealth, etc. are used as the lure to attract more talents to the supposedly more difficult jobs. The system works under the assumption that if high prestige and wealth are attached to those positions, the more talented people will be attracted to these positions. So, the basic presupposition for a status society is that human beings will put greater effort into their work mainly because there are various external rewards artificially attached to their positions. What motivates them to work is not their genuine interest in the work itself, instead, it is the external reward which they receive.

(which has nothing to do with the work itself). In other words, they do not see the work itself as meaningful. Work for them, is simply the means of achieving certain ends (which are the external rewards). The fact that so many North Americans want to be doctors or lawyers bears witness vividly to the fact that this type of incentive prevails overwhelmingly in today's capitalist societies. This type of work incentive is, as Marx sees it, one of the major causes of alienation. But it is not hard to imagine that in a different type of society, when people choose their careers, their main concern is that the type of work they intend to enter is interesting to them. In that type of society the main thing human beings will look for is meaningful work. Division of labor based on people's genuine interests and the search for meaningful work, rather than the external reward which is artificially attached with various professions, will not necessarily create class divisions and inequality. Of course, in view of the above analysis, we can see that one of the most important things involved in the change from a status society to a classless society will be a radical change of people's incentives for work. Instead of working for external reward, in a society where division of labor will not create classes, human beings will choose professions mainly because

they enjoy that type of work. If the analysis above is on the right lines, then we seem to be justified in concluding that division of labor is only a necessary condition for class division, it is not a sufficient one. Rawls sees the importance of meaningful work for a fulfilling life, but he seems not to have considered that in a society in which people's main motivation for work is their interest in the work itself, it will be highly likely that division of labor will not create class division. Rawls' model of man is, as Macpherson points out, the bourgeois model of man, i.e., a Rawlsian man is essentially a consumer instead of a doer or an exorter. He is always materially motivated. The acceptance of the difference principle confirms this observation. By accepting the difference principle, he is willing to sacrifice equality for more material gains.

If what has been said above is correct, then Rawls' claim that his theory is egalitarian becomes very dubious. Justice as Fairness tries to justify the kind of inequalities which amount to differences in life prospects for children of entrepreneurs and unskilled laborers provided that the inequalities will improve the situations of the latter. But we are still forced to ask if such society is just. The

answer must be no. Kai Nielsen puts this point nicely when he says:

It seems to me that such a society could not be a just society, let alone a perfectly just society... When people, whose only relevant difference is that one group had entrepreneurs as parents and the other had unskilled laborers as parents, have, simply because of this difference, life prospects so different that one group's entire life prospects are considerably better than the other's, then, that difference is unjust. 25

What we have to keep in mind is that the differences in the whole life prospects are not rooted in the order of nature over which we have no control, rather they are based on an artificial social system. The system can be altered so that no one will have different life prospects simply because they happen to be the son or daughter of an entrepreneur or unskilled laborer. It is hard to change one's natural talents and abilities (if such things are really natural), but it is certainly possible to make a social system in which no one has a more advantageous starting place than another. Rawls' claim that his theory is egalitarian can be accepted only if we accept the basic assumption that classlessness (in the Rawlsian sense) is impossible. But unfortunately, we have not been offered any argument to support this most crucial assumption. 26

IV. Pure Procedural Justice and Fairness

Now, we shall examine the way in which the difference principle works in distributing material goods. There are two key notions involved: first, the notion of pure procedural justice, and second, the motivation people are likely to have in a well-ordered society based on Rawls' two principles of justice. Concerning the latter, we shall only concentrate on the kind of motivation which will be generated by the difference principle.

According to Rawls, in dealing with the problem of distributive justice of material goods, pure procedural justice is the only appropriate notion we should follow. The reason for this is that when we encounter the problem of distributive shares, there is no independent criterion of what distribution is just to which we can fall back. Yet, there is a fair procedure we can adopt; and by following this procedure, any outcome that we reach will be just.

Rawls says:

In a well-ordered society, ... the distribution of material means is left to take care of itself in accordance with the idea of pure procedural justice. 27

The idea of pure procedural justice is appropriate for dealing with distributive shares when two conditions are met. Each of them is a necessary condition and they are jointly sufficient for the use of pure procedural justice. First, there is no independent criterion of what is the right result of distribution, and second, there is a correct or fair procedure such that by following it, the outcome is likewise correct or fair. The crucial thing to note is that Rawls uses the notion of pure procedural justice on two different occasions. These two different occasions embody radically different conditions such that the notion of pure procedural justice might be appropriately applied to one but definitely not to the other. The two different occasions are: first, in the derivation of the two principles of justice from the original position:

The original position is defined in such a way that it is a status quo in which any agreements reached are fair ... Thus justice as fairness is able to use the idea of pure procedural justice from the beginning. 28

Second, Rawls uses it in the actual society when the need of distributing material goods arises. After the difference principle is chosen, and people get back to their ordinary life, the distribution of material goods will be based on

the pure procedural justice in accordance with the difference principle. This is what Rawls asserts when he says that in everyday life "the distribution of material means is left to take care of itself in accordance with the idea of pure procedural justice."

The question we have to put to Rawls is, "Can both of the conditions for applying the idea of pure procedural justice be met in actual society?" Our answer is no. The condition that there is a fair procedure that can be followed, cannot be met in the real world. In order to establish a fair procedure, there must exist a state of affairs in which the parties are equally represented in every respect, and the outcome reached by following the procedure can in no way be dependent on arbitrary contingencies or on the relative balance of social forces. This is the idea behind the device of the original position. The original position is supposed to be a situation in which contractors are considered as equal and free persons in the sense that each has the right to propose any principle for consideration, and in turn each has the right to veto any principle proposed by anyone else. The veil of ignorance prevents the parties from being influenced by natural and social contingencies.

But when the idea of pure procedural justice is applied to the actual world, the conditions required for its application cannot be met. In the actual world, unlike in the original position, people are not situated in equal positions. Natural and social contingencies do play their tricks. Rawls' own conviction that those who start out as children of entrepreneurs have better life prospects than those who begin as children of unskilled laborers corroborates what we have just said. As long as class division remains in the actual world, as it does in a Rawlsian well-ordered society, pure procedural justice can never apply. In this kind of society, people are not represented equally, their whole life prospects are greatly affected by social and natural contingencies. The alleged fair procedure for distributing material wealth never existed because in order for a procedure to be fair, everyone must start equally, but the existence of class division renders this impossible. It is self-defeating on Rawls' part when he claims that Justice as Fairness attempts to mitigate the influence of natural and social contingencies while at the same time he applies the notion of pure procedural justice to the problem of distributive justice in everyday life. In applying the pure

procedural justice to the actual world, as long as class division exists, the social and natural contingencies will work their way into the final distribution. To some of Rawls' critics, it is curious that a theory of justice which starts out with egalitarian premises ends up with justifying an inequality of life prospects between members of different social classes. By pointing out that Rawls has applied the notion of pure procedural justice in the wrong place, we have been able to understand how that came about.

As to the problem of people's motivation,²⁹ it has been argued above that their main reason for accepting the inequality allowed by the difference principle is that it will improve their situations. This conviction is based on the assumption that material wealth will provide incentives to extract more effort from human beings, and hence, production will be increased. But it seems logical to infer that in a society in which material reward is the major force which makes people work, the members of the upper strata will tend to be those who are both willing and able to put in more effort for material gains. Nevertheless, as Rawls believes, the effort a person is willing and able to make³⁰ is heavily affected by his natural talents and abilities.

So, we must conclude that in using the notion of pure procedural justice on the problem of distributive justice in actual society, the goal of mitigating all natural contingencies cannot be achieved. Yet, according to Rawls, the outcome of distribution cannot be just as long as natural contingencies still play a role somewhere and have not been totally eliminated. A consistent application of the difference principle will tend to generate a certain social pattern in which the people in the upper social strata will be those endowed with certain talents and abilities, and the people in the lower social strata will be those who lack those talents and abilities. Further, as a result, people will be differentiated into various "non-comparing" groups on the bases of difference in material wealth and difference in the talents which are required for obtaining wealth. In other words, it will be a class society. In it, members of different classes will have different life prospects due to the fact that they have unequal access to the means for developing their human capacities. Children of entrepreneurs will have a much better life prospect than children of unskilled laborers. It is not without reason that some critics have argued that any society operating

on the difference principle will become a meritocratic society.³¹

Rawls is quite aware of this possibility even though he does not believe that Justice as Fairness will lead to a meritocratic society. He is also worried about the possibility that inequality in life prospects, which are justified by the difference principle, will create too much envy which will endanger the stability of a society. How does he cope with these difficulties? The way in which he deals with these problems reminds us vividly of how capitalist society works today. Rawls argues that in a well-ordered society based on his two principles, even though there exists substantial disparities of material wealth between different social classes, these disparities need not be visible, especially to the lower social strata. The very size of society makes it impossible for anyone to be frequently concerned with what is going on in the whole society. Instead people in the same social class will group together and form "plurality of associations, ... with their own secure internal life..."³² i.e., members of the same social class will find companionship in their own peer group. And the various associations based solely on material wealth more than anything

else will "tend to divide into... non-comparing groups" and between these non-comparing groups, there will be minimum communication so that the disparities between them will not attract "the kind of attention which unsettles the lives of those less well-placed."³³ This is a lively and indeed accurate description of the way in which life goes on in contemporary class society.

One cannot help but ask: Is this kind of society just? It is difficult to see how Rawls can think it possible that this kind of society is compatible with what justice requires. A well-ordered society which depends on the invisibility of the life styles between members of different classes can hardly warrant any praise from the point of view of justice. There must be something wrong with the theory of justice on which the society is based. A society can hardly be just when it relies on "the worst-off members of society continuing not to compare their position with that of the better-off."³⁴ I believe that only a deep-rooted false consciousness and constant ideological indoctrination will convince people that this kind of society can be just.

Footnotes

1. Rawls argues that the first principle has priority over the second principle, and that the second part of the second principle, i.e., the principle of equality of fair opportunity, has priority over the first part, i.e., the difference principle. The priority relation means that unless the first principle is satisfied, the second principle does not come into play. It is the same between the second part and the first part of the second principle.
2. Rawls, A Theory of Justice, (Cambridge, Massachusetts: Harvard University Press, 1971), p.546.
3. Ibid. p.151.
4. Ibid.
5. Ibid.
6. See Kai Nielsen, "Class and Justice," in Justice and Economic Distribution, (ed. by John Arthur & William H. Shaw, (Englewood Cliff, New Jersey: Prentice Hall, Inc., 1978), pp.225-245; "Radical Egalitarian Justice: Justice as Equality," Social Theory and Practice, Vol. 5, No.2 (Spring 1979), pp.209-226; "On the very Possibility of a Classless Society," Political Theory, Vol. 6, No. 2 (May, 1978), pp. 191-208; C.B. Macpherson, "Rawls' Model of Man and Society," The Philosophy of Social Sciences, Vol. 3, No. 4 (Dec.1973), pp.341-49; Brian Barry, The Liberal Theory of Justice, (Oxford: Clarendon Press, 1973), pp.50-51; David Copp, "Justice and the Difference principle," Canadian Journal of Philosophy, Vol. 4, No.2 (1974) pp.229-240; David Lyons' "Nature and Soundness of the Contract and Coherence Arguments," in Reading Rawls, Norman Daniels, ed., (New York: Basic Books Inc., 1975), pp. 141-167.
7. Rawls, "Fairness to Goodness," Philosophical Review, Vol. 84 (1975), p.543.
8. Kai Nielsen, "Radical Egalitarian Justice: Justice as Equality," p.213.

9. For the claim that inequality of wealth tends to generate inequalities of self-respect and moral autonomy, see Kai Nielsen's articles cited in footnote 6. See also Norman Daniels' article "Equal Liberty and Unequal Worth of Liberty," in Reading Rawls, pp.253-281. Daniels has forcefully criticized Rawls' distinction between "liberty" and "the worth of liberty." Externally, he argues that inequality of wealth will definitely create inequality of political liberty and right among the citizens. Internally, he argues that if it is rational for the contractors in the original position to choose equal liberty for everyone, it will be equally rational for them to choose equal worth of liberty for everyone. Furthermore, it can be shown to be irrational for them to choose the former but not the latter. For this intricate problem, see also Issaiah Berlin's article "Two Concepts of Liberty," in his book Four Essays on Liberty (Oxford: Oxford University Press, 1969), pp.118-172, and C. B. Macpherson's article "Berlin's Division of Liberty," in his book, Democratic Theory: Essays in Retrieval, (London: Oxford University Press, 1973), pp. 95-119.
10. Rawls, A Theory of Justice, p.538.
11. Lawrence Crocker, "Equality, Solidarity, and Rawls' Maximin," Philosophy and Public Affairs, Vol. 6, No. 3 (Spring 1977), pp.262-266.
12. Rawls, op. cit., p.153.
13. Ibid. p.154.
14. Ibid. p.539. Of course, mutual advantage can still be obtained through cooperation in this case. By jointly acknowledging a set of rules, they can avoid constant wars among themselves in the state of nature.
15. Robert Nozick, Anarchy, State and Utopia, (New York: Basic Books, Inc., 1974), p.184.
16. See Kai Nielsen, "Class and Justice," pp. 232-234.
17. Rawls, op. cit., p.106.
18. Ibid. p.78.
19. C. B. Macpherson, "Rawls' Model of Man," The Philosophy of Social Sciences, Vol. 3 No. 4 (Dec. 1973), p.342.

20. Ibid. p. 341-342.

21. For an illuminating discussion of the difference between Macpherson's and Rawls' notions of class division, see Kai Nielsen's article "On the Very Possibility of a Classless Society," pp. 192-195.

22. Rawls even admits at one point that socialism would be required if the stability of a well-ordered society could not be achieved in any other way. See "Fairness to Goodness," p. 546. For the compatibility between Justice as Fairness and socialism, see A Theory of Justice, pp. 270-274. However, we have to pay special attention to Rawls' claim that Justice as Fairness can also be satisfied in a capitalist society. A capitalist by owning capital, is able to extract some of the powers of the laborers for his own benefit.

23. By status society, I understand a society in which there are different social classes and people in different classes will have different life prospects. The various social statuses can be the result of a variety of reasons, e.g., inheritance, personal efforts, etc. By this definition, the Indian caste society and the contemporary capitalist society are status societies.

24. Of course, only in an affluent society in which most of the human labor is replaced by machines can this be achieved. In economically backward societies, people must spend most of their time securing the bare essentials for the maintenance of their lives, it seems almost impossible for them to adopt this attitude towards work. The Marxian principle of distribution that "from each according to his ability, and to each according to his need," can only be realized in a society, where material scarcity has largely been eliminated. However, the problem of when we can cross the threshold of material scarcity depends heavily on how much we desire. The standard of living of the low-income family in North America today would have been the dream of most of the people three hundred years ago. They would think that almost all of their desires had been satisfied if they could have all the things an average North American low-income family has today. Yet, everybody talks about

material scarcity today. Contractarians must assume that human beings are infinitely desirous, otherwise the problem of justice will not^{always} arise. The problem we face today is not so much whether all our needs can be satisfied, rather it is if the needs we have are genuine, the needs themselves have yet to be critically examined. See Herbert Marcuse, An Essay on Liberation, (Boston: Beacon Press, 1969).

25. Kai Nielsen, "Class and Justice," pp.229-230.
26. For a conception of justice which is more egalitarian than Rawls', see Kai Nielsen, "Radical Egalitarian Justice: Justice as Equality," pp.211-212. He proposes two principles of justice as follows:
 1. Each person is to have an equal right to the most extensive total system of equal basic liberties and opportunities (including equal opportunities for meaningful work, for self-determination and political participation) compatible with a similar treatment of all. (This principle gives expression to a commitment to attain and / or sustain equal moral autonomy and equal self-respect).
 2. After provisions are made for common social (community) values, for capital overhead to preserve the society's productive capacity and allowances are made for differing unmanipulated needs and preferences, the income and wealth (the common stock of means) is to be so divided that each person will have a right to an equal share. The necessary burden requisite to enhance well-being is also to be equally shared, subject, of course, to limitations by differing abilities and differing situations (natural environment, not class position).
27. Rawls, A Theory of Justice, p.545.
28. Ibid. p.120.
29. We have argued earlier that to conceive all human relationships, as contractual is incompatible with the view that human beings will undergo fundamental changes upon entering society; the reasons for entering the contract will also be the reasons for maintaining the society created by the contract. So, the contractors in the original position and people in the ordinary world will share the same motivation, i.e., self-interestedness.

30. In criticizing the precept that distributive justice should be in accordance with a person's moral worth, Rawls argues against the principle of distribution according to effort. But he does not realize the criticism can be directed against himself with the same force. See A Theory of Justice, p.312.

31. Thomas Nagel, "Rawls on Justice," in Reading Rawls, ed. by Norman Daniels, pp.12-13, The footnote.

32. Rawls, A Theory of Justice, p.536.

33. Ibid. p.536-537.

34. Russell Keat and David Miller, "Understanding Justice," Political Theory, Vol. 2, No. 1 (1974) p.24.

Concluding Remarks

In the last decade or so, philosophers in the English-speaking world have shown great interest in the problem of justice. Moral and political philosophy after a long period of dormancy, have once again become lively areas for philosophizing. In the center of the revitalized moral and political discourses stands the theory of justice. Without any doubt, the enthusiasm is aroused in part by the appearance of Rawls' A Theory of Justice, which is a systematic treatise dealing with the problems of distributive justice. But more importantly, the great political and social upheavals in the 1960's in North America and elsewhere, undoubtedly have prompted people to look into the basic designs of society to find the roots of these troubles. The philosophical speculations about the problems of justice, including Rawls' work, can be viewed as philosophical responses to the experience of turmoil people had. Marshall Cohen in his review of A Theory of Justice notes:

All great political philosophies of the past -- Plato's, Hobbes', Rousseau's -- have responded to the realities of contemporary politics, and it is therefore not surprising that Rawls' pene-

trating account of the principles to which our public life is committed should appear at a time when these principles are persistently being obscured and betrayed.

Cohen's statement, of course, can be applied with equal force to moral philosophies.

The upheaval in the sixties in the U. S. and some other Western societies -- the Civil Rights and Black Liberation movements, the Hippie movement, the Anti-Vietnam War movements, and the student-worker rebellion of 1968 in France, brought millions of people into conflict with existing political institutions and policies. It also made millions of people doubt the justice of the basic structure of the society they lived in. The optimistic picture of Western societies in general, and American society in particular, as affluent and just societies with only a few residual flaws to be ironed out in the not too distant future by more advanced technology is no longer naively believed. The sceptics called into question not only the integrity and competence of specific leaders or the justice of particular laws and policies, they suspected the justice of the basic structure of society itself. They raised fundamental questions about the justice of basic political, social and economic designs which affect greatly everyone's opportunities or life pros-

pects in society. We can safely say that Rawls, who worked his book throughout this period, must have been concerned with the issues raised by some of the sceptics (as his penetrating discussion of the problems of civil disobedience attests). Furthermore, we may even say that Rawls' decision to take the basic structure of society as his subject matter in dealing with the problems of justice is not completely unaffected by what was going on in the sixties.

The basic structure of Western society, of course, embodies many of the utilitarian ideals. For a long period of time, utilitarianism has been the predominant moral theory in the English-speaking tradition. Social and economic institutions are set up in accordance with the principle of utility to maximize the total utility. The increment of Gross National Product has been the most important goal of society. People are supposed to work for "the greatest happiness of the greatest number" of their fellow men; minorities should submit to the interests of the majority. According to utilitarianism, an action is right if it results in a net positive increment to the aggregate of individual utilities, even though such action may render some individuals worse off. In an individualistic society in

which everybody cares only about his own interest, the possible inequality caused by the goal of maximizing total utility will result in bitter feeling on the part of the people whose interests have been sacrificed. In order for a society to function smoothly, the utilitarians must presuppose that people are benevolent; they cannot be individualists. Contrary to common belief, Rawls and Derek Parfit have argued convincingly that utilitarianism is not individualistic; it requires that the participants in a social scheme² be benevolent.

The requirement that everyone be benevolent is apparently incompatible with the predominant ideology, i.e., individualism, in capitalist society today. In an individualistic society, everyone is to maximize his own utility. The feeling of benevolence is obviously lacking. There exists a fundamental conflict between the two rival philosophical doctrines prevailing in Western society today. On the one hand, we have utilitarianism, in accordance with it major social and economic institutions are set up. Utilitarianism presupposes that people must be benevolent and non-individualistic. On the other hand, individualism is the prevailing ideology everybody subscribes to. This fundamental conflict can

be resolved only by a philosophical overhaul. Rawls' main task is to resolve this fundamental conflict.

Utilitarianism is unacceptable to Rawls because it sometimes may require that some people sacrifice their interest for the welfare of others. Justice as fairness, as I see it, is an attempt to provide a philosophical foundation for an individualistic society which is the reality we live with today. Despite Rawls' belief that his theory will provide a base for social union, I have argued extensively in the last two chapters that contractarianism will inevitably lead to a civil society. The gist of Rawls' theory can be summarized in his own words:

Each person possesses an inviolability, founded on justice that even the welfare of society as a whole cannot override ... It does not allow that the sacrifices imposed on a few are outweighed by the larger sum of advantages enjoyed by many. 3

Rawls might be wrong, but he is perfectly relevant. His theory provides a challenge not only to the utilitarians, but also to anyone in the future who attempts to deal with seriously the problems involved in moral, social, and political philosophy.

Footnotes

1. Marshall Cohen, "Social Contract Explained and Defended," New York Times Book Review, 16 July 1972, p.1.
2. Rawls, A Theory of Justice (Cambridge, Massachusetts: Harvard University Press, 1971), pp.27-33. Derek Parfit, "Later Selves and Moral Principles," in Philosophy and Personal Relations, ed. by Alan Montefiore (Montreal: McGill-Queen's University Press, 1973), pp.137-69.
3. Rawls, op. cit. pp.3-4.

Bibliography

Rawls' Works:

- Rawls, John, "Outline of a Decision Procedure for Ethics," The Philosophical Review, 60 (1951), pp.177-97.
- , "Two Concepts of Rules," The Philosophical Review, 64 (1955), pp.3-32.
- , "Justice as Fairness," The Philosophical Review, 67 (1958), pp.164-94.
- , "The Sense of Justice," The Philosophical Review, 72 (1963), pp.281-305.
- , "Distributive Justice," in Philosophy, Politics and Society, Third Series, ed. by Peter Laslett and W.G. Runciman (London: Basil Blackwell Ltd., 1967).
- , "Justice as Reciprocity," in Utilitarianism: John Stuart Mill, with Critical Essays, ed. by Samuel Gorovitz (New York: Bobbs-Merrill Co. Inc., 1971).
- , 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), pp.556-7.
- , "Some Reasons for the Maximin Criterion," a symposium on Concepts of Distributional Equity, The American Economic Review, 64 (1974), pp.141-46.
- , "Reply to Alexander and Musgrave," The Quarterly Journal of Economics, 88 (1974), pp.633-55.
- , "Fairness to Goodness," The Philosophical Review, 84, (1974) pp.536-54.

-----, "The Independence of Moral Theory," Presidential Address to the American Philosophical Association, Eastern Division, 1974, in Proceedings and Addresses of the American Philosophical Association, 48 (1975), pp.5-22.

Works on Rawls and Related Works:

Barber, Benjamin R., "Justifying Justice: Problems of Psychology, Measurement, and Politics in Rawls," The American Political Science Review, 69 (1975), pp.663-74.

Barry, Brian, "On Social Justice," The Oxford Review, (Trinity Term, 1967), pp.33-43.

-----, The Liberal Theory of Justice, (London: Oxford University Press, 1973).

Bates, Stanley, "The Motivation to be Just," Ethics, 85 (1974), pp.1-17.

Benn, S.I., and Peters, R.S., Social Principles and the Democratic State (London: George Allen and Unwin Ltd., 1959).

Berlin, Isaiah, Four Essays on Liberty (Oxford: Oxford University Press, 1969).

Campbell, T.D., "Humanity before Justice," The British Journal of Political Science, 4 (1974), pp.1-16.

Care, Norman S., "Contractualism and Moral Criticism," The Review of Metaphysics, 23 (1969), pp.85-101.

Carr, Spencer, "Rawls, Contractarianism, and our Moral Intuitions," The Personalist, 56 (1975), pp.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), pp.147-69.

Crocker, Lawrence, "Equality, Solidarity, and Rawls' Maximin,"
Philosophy and Public Affairs, Vol. 6, No.3
(Spring 1977) pp.262-66.

Cunningham, R.L., "Justice: Efficiency or Fairness?" The Personalist, 52 (1971) pp.253-81.

Daniels, Norman, (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), pp.253-81.

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), pp.124-40.

Dworkin, Ronald, "The Origin Position," in Reading Rawls: Critical Studies of a Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), pp.16-52.

Elle, Larry, "Individualism and Class Consciousness," Radical Philosopher's Newsjournal, IX (Fall 1977), pp.1-13.

Eshete, Andreas, "Contractarianism and the Scope of Justice," Ethics, 85 (1974) pp.38-49.

Feinberg, Joel, "Rawls and Intuitionism," in Reading Rawls: Critical Studies of a Theory of Justice, ed. by Norman Daniels (New York: Basic Books Inc., 1975), pp.108-23.

Fishkin, James, "Justice and Rationality: Some Objections to the Central Argument in Rawls' Theory," The American Political Science Review, 69 (1975), pp.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), pp.53-80.

- Frankel, Charles, "The New Egalitarianism and the Old,"
Commentary, (September, 1979), pp.54-61.
- Gauthier, David, "Justice and Natural Endowment: Toward a
Critique of Rawls' Ideological Framework,"
Social Theory and Practice, 3 (1974), pp.3-26.
- Gauthier, David, "Why Ought One Obey God: Reflections on
Hobbes and Locke," Canadian Journal of Philosophy,
Vol. 7, No.3 (1977), pp.425-66.
- , "Social Contract as Ideology," Philosophy
and Public Affairs, Vol. 6, No.2 (W.1977), pp.130-64.
- Hare, R.M., "Rawls' Theory of Justice," in Reading Rawls:
Critical Studies of a Theory of Justice, ed.
by Norman Daniels (New York: Basic Books Inc.,
1975), pp.81-107.
- Honderich, Ted, "The Use of the Basic Proposition of a Theory
of Justice," Mind, 84 (1975), pp.63-78.
- Hospers, John, "A Review of Rawls' A Theory of Justice,"
The Personalist, 55 (1974), pp.71-77.
- Hume, David, An Inquiry Concerning the Principles of Morals,
ed. with an Introduction by Charles W. Hendel
(New York: The Bobbs-Merrill Company Inc., 1957).
- Kant, Immanuel, Foundations of the Metaphysics of Morals,
(New York: The Bobbs-Merrill Company, Inc., 1959).
- Keat, Russell, and Miller, David, "Understanding Justice,"
Political Theory, Vol.2, No.1 (1974), pp.3-31.
- Keyt, David, "The Social Contract as an Analytic, Justifica-
tory and Polemic Device," The Canadian Journal
of Philosophy, 4 (1974), pp.241-52.
- Kleinberger, Aharon F., "The Social-Contract Strategy for
the Justification of Moral Principles," Journal
of Moral Education, Vol.5, No.2 (1978), pp.107-26.
- Kuhn, Thomas, The Structure of Scientific Revolution (Chicago:
The University of Chicago Press, 1962).

- Lakatos, Imre, and Musgrave, Alan (ed.), Criticism and the Growth of Knowledge (Cambridge: Cambridge University Press, 1970).
- Lessnoff, Michael, "John Rawls' Theory of Justice," Political Studies, 19 (1971), pp.63-80.
- Levine, Andrew, "Rawls' Kantianism," Social Theory and Practice, 3 (1973), pp.47-63.
- Lukes, Steven, "Relativism: Cognitive and Moral," The Aristotelian Society 48 (1974), pp.165-89.
- , "An Archimedean Point," Observer Review, (4 June, 1972).
- , Individualism (Oxford: Basil Blackwell, 1973).
- Lyons, David, "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), pp.141-67.
- Macpherson, C.B., "Rawls' Model of Man and Society," Philosophy of the Social Sciences, 3 (1973), pp.341-7.
- , The Political Theory of Possessive Individualism (London: Oxford University Press, 1962).
- , "Class, Classlessness, and the Critique of Rawls: A reply to Nielsen," Political Theory, Vol.6, No.2 (May, 1978), pp.209-211.
- Marcuse, Herbert, One-Dimension Man, (Boston: Beacon Press, 1964).
- , An Essay on Liberation, (Boston: Beacon Press, 1969).
- Marshall, J., "The Failure of Contract as Justification," Social Theory and Practice, 3 (1975), pp.441-59.
- Marx, Karl, Early Writings, tr. and ed. by T.B. Bottomore (New York: McGraw-Hill Book Company, 1964).

- Marx, Karl, and Engels, Frederick, The German Ideology, Part One, with Selections from Part Two and Three and Supplementary Texts, ed. with Introduction by C. J. Arthur, (New York: International Publishers, 1970).
- Mcbride, W., "Social Theory Sub Specie Aeternitatis: A New Perspective," The Yale Law Journal, 5(1972), pp.980-1003.
- , "The Concept of Justice in Marx, Engel, and Others," Ethics, 85 (1975), pp.204-218.
- Miller, Richard, "Rawls and Marxism," Philosophy and Public Affairs, 3 (1974), pp.167-94.
- Nagel, Thomas, "Rawls on Justice," The Philosophical Review, 82 (1973), pp.220-34.
- Nielsen, Kai, "A Note on Rationality," The Journal of Critical Analysis, 4(1972), pp.16-19.
- , "On Philosophic Method," International Philosophical Quarterly, 16 (1976), pp.349-368.
- , "Rawls and Classist Amoralism," Mind, 86 (1977), pp.19-30.
- , "Rawls' Defense of Morality: Rationality, Amoralism and the Problem of Congruence," The Personalist, Vol.59, No. 1 (1978), pp.93-100.
- , "Class and Justice," in Justice and Economic Distribution, ed. by John Arthur and William H. Shaw (Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1978), pp. 225-45.
- , "On the Diversity of Moral Beliefs," Cultural Hermeneutics, 2 (1974), pp.281-303.
- , "On the Very Possibility of a Classless Society," Political Theory, Vol.6. No.2 (May,1978), pp.191-208.
- , "Radical Egalitarian Justice? Justice as Equality," Social Theory and Practice, Vol.5, No.2 (Spring 1979), pp.209-226.

Nielsen, Kai, and Shiner, Roger, New Essays on Contract Theory, (Guelph, Ontario: Canadian Association for Publishing in Philosophy, 1977).

Nowell-Smith, P.H., "A Theory of Justice," Philosophy of the Social Science, 3 (1973), pp.315-39.

Nozick, Robert, "Distributive Justice," Philosophy and Public Affairs, 3 (1973), pp.45-126.

-----, Anarchy, State, and Utopia, (New York: Basic Books Inc., 1974).

Parfit, Derek, "Later Selves and Moral Principles," in Philosophy and Personal Relations, ed. by Alan Montefiore (Montreal: McGill-Queen's University Press, 1973), pp.137-69.

Popper, Karl, Conjectures and Refutations, (New York: Harper & Row, 1963).

Prichard, H.A., Moral Obligation and Duty and Interest (London: Oxford University Press, 1968).

Proudfoot, Wayne, "Rawls on the Individual and the Social," The Journal of Religious Ethics, 2/2 (1974), pp.107-28.

Rousseau, Jean Jacques, The Social Contract, Translated with Introduction by G.D.H. Cole (London: J.M. Dent and Sons Ltd., 1913).

Schwartz, Adina, "Moral Neutrality and Primary Goods," Ethics, 83 (1973), pp.294-307.

Searle, John, "What is a Speech Act?" in Philosophy in America, ed. by Max Black (London: George Allen and Unwin Ltd., 1965), pp.221-39.

Singer, Marcus, "The Method of Justice: Reflections on Rawls," The Journal of Value Inquiry, Vol.X, No.4 (W1976), pp.286-316.

Singer, Peter, "Sidgwick and Reflective Equilibrium," The Monist, 58 (1974), pp.490-517.

Snare, Frank, "John Rawls and the Methods of Ethics," Philosophy and Phenomenological Research, 36 (1975), pp.100-12.

Teitelman, Michael, "The Limits of Individualism," in "Symposium: A Theory of Justice," The Journal of Philosophy, 69 (October 5, 1972), pp.545-56.

Thomas, G.B., "On Choosing a Morality," Canadian Journal of Philosophy, 52(1975), pp.357-74,

Urmson, J.O., "A Defence of Intuitionism," The Aristotelian Society, 75 (1974-75), pp.111-19.

Wolfe, Robert Paul, In Defense of Anarchism (New York: Harper & Row, 1970).

-----, Understanding Rawls (Princeton, New Jersey: Princeton University Press, 1977).

Appendix

Abstract of The Limits of Contractarianism: Rawls' Moral Methodology and Ideological Framework

The publication of Rawls' A Theory of Justice in 1971 is chiefly responsible for the revival of ethics and political philosophy since the beginning of the 1970's in the Anglo-American philosophical world.

Rawls calls his theory Justice as Fairness. It is a theory which attempts to revitalize the contractarian tradition of moral philosophy represented by Locke, Kant, and Rousseau.

The aim of this dissertation is to provide a critical examination of Rawls' theory of justice from two perspectives. First, we have examined the methods Rawls uses in deriving and justifying the two principles of his theory; second, we have tried to unpack the ideological framework embodied in Justice as Fairness.

There are two methods Rawls employs in constructing his theory. First, the contractarian method, and second, the method of reflective equilibrium. With regard to the contractarian method, we have argued that Rawls' claim that moral principles are the objects of rational choice in the original position is problematic. Rawls' arguments for this

claim have been critically examined and shown to be defective. As to the method of reflective equilibrium, we have argued that it does not really carry us beyond relativism and intuitionism.

Concerning the issues of Rawls' ideological framework, we have argued that the adoption of the contractarian approach has certain ideological implications. The contractarian method is not neutral, and contractarianism presupposes some substantive philosophical theses. In particular, it presupposes individualism. We have also critically examined the difference principle and shown that the adoption of this principle will lead to a class-divided society.