Chapter I. Justice as Fairness

1. The Role of Justice

Justice is the first virtue of social institutions, as truth is of systems of thought. A theory however elegant and economical must be rejected or revised if it is untrue; likewise laws and institutions no matter how efficient and well arranged must be reformed or abolished if they are unjust. Each person possesses an inviolability founded on justice that even the welfare of society as a whole cannot override. For this reason justice denies that the loss of freedom for some is made right by a greater good shared by others. It does not allow that the sacrifices imposed on a few are outweighed by the larger sum of advantages enjoyed by many. Therefore in a just society, the liberties of equal citizenship are taken as settled; the rights secured by justice are not subject to political bargaining or to the calculus of social interests. The only thing that permits us to acquiesce in an erroneous theory is the lack of a better one; analogously, an injustice is tolerable only when it is necessary to avoid an even greater injustice. Being first virtues of human activities, truth and justice are uncompromising.

These propositions seem to express our intuitive conviction of the primacy of justice. No doubt they are expressed too strongly. In any event I wish to inquire whether these contentions or others similar to them are sound, and if so how they can be accounted for. To this end it is necessary to work out a theory of justice in the light of which these assertions can be interpreted and assessed. I shall begin by considering the role of the principles of justice. Let us assume, to fix ideas, that a society is a more or less self-sufficient association of persons who in their relations to one another recognize certain rules of conduct as binding and who for the most part act in accordance with them. Suppose further that these rules specify a system of cooperation designed to advance the good of those taking part in it. Then, although a society is a cooperative venture for mutual advantage, it is typically marked by a conflict as well as by an identity of interests. There is an identity of interests since social cooperation makes possible a better life for all than any would have if each were to live solely by his own efforts. There is a conflict of interests since persons are not indifferent as to how the greater benefits produced by their collaboration are distributed, for in order to pursue their ends they each prefer a larger to a lesser share. 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 social 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.

Now let us say that a society is well ordered when it is not only designed to advance the good of its members but when it is also effectively regulated by a public conception of justice. That is, it is a society in which (1) everyone accepts and knows that the others accept the same principles of justice, and (2) the basic social institutions generally satisfy and are generally known to satisfy these principles. In this case while men may put forth excessive demands on one another, they nevertheless acknowledge a common point of view from which their claims may be adjudicated. If men's inclination to self-interest makes their vigilance against one another necessary, their public sense of justice makes their secure association together possible. Among individuals with disparate aims and purposes a shared conception of justice establishes the bonds of civic friendship; the general desire for justice limits the pursuit of other ends. One may think of a public conception of justice as constituting the fundamental charter of a well-ordered human association.
3. The Main Idea of the Theory of Justice

My aim 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. In order to do this we are not to think of the original contract as one to enter into 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. They 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. These principles are to regulate all further agreements; they specify the kinds of social cooperation that can be entered into and the forms of government that can be established. This way of regarding the principles of justice I shall call justice as fairness.

Thus we are to imagine that those who engage in social cooperation choose together, in one joint act, the principles which are to assign basic rights and duties and to determine the division of social benefits. Men are to decide in advance how they are to regulate their claims against one another and what is to be the foundation charter of their society. Just as each person must decide by rational reflection what constitutes his good, that is, the system of ends which it is rational for him to pursue, so a group of persons must decide once and for all what is to count among them as just and unjust. The choice which rational men would make in this hypothetical situation of equal liberty, assuming for the present that this choice problem has a solution, determines the principles of justice.

In justice as fairness the original position of equality corresponds to the state of nature in the traditional theory of the social contract. This original position is not, of course, thought of as an actual historical state of affairs, much less as a primitive condition of nature. It is understood as a purely hypothetical situation characterized so as to lead to a certain conception of justice. Among the essential features of this situation is that no one knows his place in society, his class position or social status, nor does any one know his fortune in the distributions 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. This ensures that no one is advantaged or disadvantaged in the choice of principles by the outcome of natural chance or the contingency of social circumstances. Since all are similarly situated and no one is able to design principles to favor his particular condition, the principles of justice are the result of a fair agreement or bargains. For given the circumstances of the original position, the symmetry of everyone's relations to each other, this initial situation is fair between individuals as moral persons, that is, as rational beings with their own ends and capable, I shall assume, of a sense of justice. The original position is, one might say, the appropriate initial status quo, and thus the fundamental agreement reached in it is fair. This explains the propriety of the name "justice as fairness"; it conveys the idea that the principles of justice are agreed to in an initial situation that is fair. The name does not mean that the concepts of justice and fairness are the same, any more than the phrase "poetry as metaphor" means that the concepts of poetry and metaphor are the same.

Justice as fairness begins, as I have said, with one of the most general of all choices which persons might make together, namely, with the choice of the first principles of a conception of justice which is to regulate all subsequent criticism and reform of institutions. Then, having chosen a conception of justice, we can suppose that they are to choose a constitution and a legislature to enact laws, and so on, in accordance with the principles of justice initially agreed upon. Our social situation is just if it is such that by this sequence of hypothetical agreements we have contracted into the general system of rules which defines it. Moreover, assuming that the original position does determine a set of principles (that is, that a particular conception of justice would be chosen), it will then be true that whenever social institutions satisfy these principles those engaged in them can say to one another that they are cooperating on terms to which they would agree if they were free and equal persons whose relations with respect to one another were fair. They could all view their arrangements as meeting the stipulations which they would acknowledge in an initial situation that embodies widely accepted and reasonable constraints on the choice of principles. The general recognition of this fact would provide the basis for a public acceptance of the corresponding principles of justice. No society can, of course, be a scheme of cooperation which men enter voluntarily in a literal sense; each person finds himself placed at birth in some particular position in some particular society, and the nature of this position materially affects his life prospects. Yet a society satisfying the principles of justice as fairness comes as close as a society can to being a voluntary scheme, for it meets the principles which free and equal persons would assent to under circumstances that are fair. In this sense its members are autonomous and the obligations they recognize self-imposed.

One feature of justice as fairness is to think of the parties in the initial situation as rational and mutually disinterested. This does not mean that the parties are egoists, that is, individuals with only certain kinds of interests, say, in wealth, prestige, and domination. But they are conceived as not taking an interest in one another's interests. They are to presume that even their spiritual aims may be opposed, in the way that the aims of those of different religions may be opposed. Moreover, the concept of rationality must be interpreted as far as possible in the narrow sense, standard in economic theory, of taking the most effective means to given ends. . . . One must try to avoid introducing into it any controversial ethical elements. The initial situation must be characterized by stipulations that are widely accepted.

In working out the conception of justice as fairness one main task clearly is to determine which principles of justice would be chosen in the original position. To do this we must describe this situation in some detail and formulate with care the problem of choice which it presents . . . . It may be observed, however, that once the principles of justice are thought of as arising from an original agreement in a situation of equality, it is an open question whether the principle of utility would be acknowledged. Offhand it hardly seems likely that persons who view themselves as equals, entitled to press their claims upon one another, would agree to a principle which may require lesser life prospects for some simply for the sake of a greater sum of advantages enjoyed by others. Since each desires to protect his interests, his capacity to advance his conception of the good, no one has a reason to acquiesce in an enduring less for himself in order to bring about a greater net balance of satisfaction. In the absence of strong and lasting benevolent impulses, a rational man would not accept a basic structure merely because it maximized the algebraic sum of advantages irrespective of its permanent effects on his own basic rights and interests. Thus it seems that the principle of utility is incompatible with the conception of social cooperation among equals for mutual advantage. It appears to be inconsistent with the idea of reciprocity implicit in the notion of a well-ordered society. Or, at any rate, so I shall argue.

I shall maintain instead that the persons in the initial situation would choose two rather different principles: the first requires equality in the assignment of basic rights and duties, while the second holds that social and economic inequalities, for example inequalities of wealth and authority, are just only if they result in compensating benefits for everyone, and in particular for the least advantaged members of society. These principles rule out justifying institutions on the grounds that the hardships of some are offset by a greater good in the aggregate. It may be expedient but it is not just that some have less in order that others may prosper. But there is no injustice in the greater benefits carried by a few provided that the situation of persons not so fortunate is thereby improved. The intuitive idea is that since everyone's well-being depends upon a scheme of cooperation without which no one could have a satisfactory life, the division of advantages should be such as to draw forth the willing cooperation of everyone taking part in it, including those less well situated. Yet this can be expected only if reasonable terms are proposed. The two principles mentioned seem to be a fair agreement on the basis of which those better endowed, or more fortunate in their social position, neither of which we can be said to deserve, could expect the willing cooperation of others when some workable scheme is a necessary condition of the welfare of all. Once we decide to look for a conception of justice that nullifies the accidents of natural endowment and the contingencies
of social circumstance as counters in quest for political and economic advantage, we are led to these principles. They express the result of leaving aside those aspects of the social world that seem arbitrary from a moral point of view.

4. The Original Position and Justification

I have said that the original position is the appropriate initial status quo which insures that the fundamental agreements reached in it are fair. This fact yields the name "justice as fairness." It is clear, then, that I want to say that one conception of justice is more reasonable than another, or justifiable with respect to it, if rational persons in the initial situation would choose its principles over those of the other for the role of justice.

Conceptions of justice are to be ranked by their acceptability to persons so situated. Understood in this way the question of justification is settled by working out a methodology of deliberation; we have to ascertain which principles it would be rational to adopt given the contractual situation. This connects the theory of justice with the theory of rational choice.

If this view of the problem of justification is to succeed, we must, of course, describe in some detail the nature of this choice problem. A problem of rational adjudication has a definite answer only if we know the beliefs and interests of the parties, their relations with respect to one another, the alternatives between which they are to choose, the procedure whereby they make up their minds, and so on. As the circumstances are presented in different ways, correspondingly different principles are accepted. The concept of the original position, as I shall refer to it, is that of the most philosophically favored interpretation of this initial choice situation for the purposes of a theory of justice.

But how are we to decide what is the most favored interpretation? I assume, for one thing, that there is a broad measure of agreement that principles of justice should be chosen under certain conditions. To justify a particular description of the initial situation one shows that it incorporates these commonly shared presumptions. One argues from widely accepted but weak premises to more specific conclusions. Each of the presumptions should by itself be natural and plausible; some of them may seem innocent or even trivial. The aim of the contract approach is to establish that taken together they impose significant bounds on acceptable principles of justice. The ideal outcome would be that these conditions determine a unique set of principles; but I shall be satisfied if they suffice to rank the many traditional conceptions of social justice.

One should not be misled, then, by the somewhat unusual conditions which characterize the original position. The idea here is simply to make vivid to ourselves the restrictions that it seems reasonable to impose on arguments for principles of justice, and therefore on these principles themselves. Thus it seems reasonable and generally acceptable that no one should be advantaged or disadvantaged by natural fortune or social circumstances in the choice of principles. It also seems widely agreed that it should be impossible to tailor principles to the circumstances of one's own case. We should insure further that particular inclinations and aspirations, and persons' conceptions of their good do not affect the principles adopted. The aim is to rule out those principles that it would be rational to propose for acceptance, however the chance of success, only if one knew certain things that are irrelevant from the standpoint of justice. For example, if a man knew that he was wealthy, he might find it rational to advance the principle that various taxes for welfare measures be counted unjust; if he knew that he was poor, he would most likely propose the contrary principle. To represent the desired restrictions one imagines a situation in which everyone is deprived of this sort of information. One excludes the knowledge of those contingencies which sets men at odds and allows them to be guided by their prejudices. In this manner the veil of ignorance is arrived at in a natural way. This concept should cause no difficulty if we keep in mind the constraints on arguments that it is meant to express. At any time we can enter the original position, so to speak, simply by following a certain procedure, namely, by arguing for principles of justice in accordance with these restrictions.

It seems reasonable to suppose that the parties in the original position are equal. That is, all have the same rights in the procedure for choosing principles; each can make proposals, submit reasons for their acceptance, and so on. Obviously the purpose of these conditions is to represent equality between human beings as moral persons, as creatures having a conception of their good and capable of a sense of justice. The basis of equality is taken to be similarity in these two respects. Systems of ends are not ranked in value; and each man is presumed to have the requisite ability to understand and to act upon whatever principles are adopted. Together with the veil of ignorance, these conditions define the principles of justice as those which rational persons concerned to advance their interests would consent to as equals when none are known to be advantages or disadvantaged by social and natural contingencies.

There is, however, another side to justifying a particular description of the original position. This is to see if the principles which would be chosen match our considered convictions of justice or extend them in an acceptable way. We can note without applying these principles would lead us to make the same judgments about the basic structure of society which we now make intuitively and in which we have the greatest confidence; or whether, in cases where our present judgments are in doubt and given with hesitancy, one would be better off without which one can affirm on reflection. There are questions which we feel sure must be answered in a certain way. For example, we are confident that religious intolerance and racial discrimination are unjust. We think that we can examine these things with care and have reached what we believe is an impartial judgment not likely to be distorted by an excessive attention to our own interests. These convictions are provisional fixed points on which we assume any conception of justice must fit. But we have much less assurance as to what is the correct distribution of wealth and authority. Here we may be looking for a way to remove our doubts. We can check an interpretation of the initial situation, then, by the capacity of its principles to accommodate our firmest convictions and to provide guidance where guidance is needed.

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, 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 fully refined and adjusted.

This state of affairs I refer to as reflective equilibrium. 1 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. At the moment nothing is in order. But this equilibrium is not necessarily stable. It is liable to be upset by further examination of the conditions which should be imposed on the contractual situation and by particular cases which may lead us to revise our judgments. Yet for the time being it represents what we can regard as coherent and to justify our convictions of social justice. We have reached a conception of the original position.

I shall not, of course, actually work through this process. Still, we may think of the interpretation of the original position that I shall present as the result of such a hypothetical course of reflection. It represents the attempt to accommodate within one scheme both reasonable philosophical conditions on principles as well as our considered judgments of justice. In arriving at the favored interpretation of the initial situation there is no point at which an appeal is made to self-evidence in the traditional sense either of general conceptions or particular convictions. I do not claim for the principles of justice proposed that they are necessary truths or derivable from such truths. A conception of justice cannot be deduced from self-evident premises or conditions or principles; instead, its justification is a matter of the mutual support of many considerations, of everything fitting together into one coherent view.

A final comment. 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
It is essential to keep in mind that in a teleological theory the good is defined independently from the right. This means two things. First, the theory accounts for our considered judgments as to what things are good for which judgments of value as a separate class of judgments intuitively distinguishable by common sense, and then suggests that the right is maximizing the good as already specified. Second, the theory enables one to judge the goodness of things without referring to what is right. For example, if pleasure is to be the sole good, then presumably pleasures can be recognized and ranked in value by criteria that do not presuppose any standards of right, or what we would normally think of as such. Whereas if the distribution of goods is also counted as a good, perhaps a higher order one, and the theory directs us to produce more goods (including the good of distribution among others), we no longer have a teleological view of the classical sense. The problem of distribution falls under the concept of right as one intuitively understands it, and so the theory lacks an independent definition of the good. The clarity and simplicity of classical teleological theories derives largely from the fact that they are more unambiguously specified class; the one being characterized separately while the other is then connected with it by a maximizing principle.

Tteleological doctrines differ, pretty clearly, according to how the conception of the good is specified. If it is taken as the realization of human excellence in the various forms of culture, we have what may be called perfectionism. This notion is found in Aristotle and Nietzsche, among others. If the goal is defined as pleasure, we have hedonism, or as happiness, eudaimonism, and so on. I shall understand the principle of utility in its classical form as defining the good as the satisfaction of desire, or perhaps better, as the satisfaction of rational desire. This accords with the view in all essential ways, and provides, I believe, a fair interpretation of it. The appropriate terms of social cooperation are settled by whatever in the circumstances will achieve the greatest sum of satisfaction of the rational desires of individuals. It is impossible to deny the initial plausibility and attractiveness of this conception.

The striking feature of the utilitarian view of justice is that it does not matter, except indirectly, how this sum of satisfaction is distributed among individuals any more than it matters, except indirectly, how one man distributes his satisfactions over time. The correct distribution in either case is that which yields the maximum fulfillment. Society must allocate its means of satisfaction whatever these are, rights and duties, opportunities and privileges, and various forms of wealth, so as to achieve this maximum if it can. But in itself no distribution of satisfaction is better than another except that the more equal distribution is to be preferred to break ties. It is true that certain common-sense precepts of justice, particularly those which concern the protection of liberty and rights, or which express the claims of desert, seem to contradict this contention. But from a utilitarian standpoint the explanation of these precepts and of their seemingly stringent character is that they are these precepts which experience shows should be strictly respected and departed from only under exceptional circumstances if the sum of advantages is to be maximized. Yet, as with all other precepts, those of justice are derivative from the one end of attaining the greatest balance of satisfaction. Thus there is no reason in principle why the greater gains of some should not compensate for the lesser losses of others; or more concretely, why the violation of the liberty of a few might not be justified by the greater good shared by many. It simply happens that under most conditions, at least in a reasonably advanced stage of civilization, the greatest sum of advantages is not attained in this way. No doubt the strictness of common-sense precepts of justice has a certain usefulness in limiting men's propensities to injustice and to socially injurious actions, but the utilitarian believes that to affirm this strictness as a first principle of morals is a mistake. For just as it is rational for one man to maximize the fulfillment of his system of desires, it is right for a society to maximize the net balance of satisfaction taken over all of its members.

The most natural way, then, of arriving at utilitarianism (although not, of course, the only way of doing so) is to adopt for society as a whole the principle of rational choice for one man. Once this is recognized, the place of the impartial spectator and the emphasis on sympathy in the history of utilitarian thought is readily understood. For it is by the conception of the impartial spectator and the use of sympathetic identification in guiding our imagination that the principle for one man is applied to society. It is this spectator
who is conceived as carrying out the required organization of the desires of all persons into a coherent system of desire; it is by this construction that many persons are fused into one. Endowed with ideal powers of sympathy and imagination, the impartial spectator is the perfectly rational individual who identifies with and experiences the desires of others as if these desires were his own. In this way he ascertains the intensity of these desires and assigns them their appropriate weight in the one system of desire the satisfaction of which the ideal legislator then tries to maximize by adjusting the rules of the social system. On this conception of society separate individuals are thought of as so many different lines along which rights and duties are to be assigned and scarce means of satisfaction allocated in accordance with rules so as to give the greatest fulfillment of wants. The nature of the decision made by the ideal legislator is not, therefore, materially different from that of an entrepreneur deciding how to maximize his profit by producing this or that commodity, or that of a consumer deciding how to maximize his satisfaction by the purchase of this or that collection of goods. In each case there is a single person whose system of desires determines the best allocation of limited means. The correct decision is essentially a question of efficient administration. This view of social cooperation is the consequence of extending to society the principle of choice for one man, and then, to make this extension work, conflating all persons into one through the imaginative acts of the impartial sympathetic spectator. Utilitarianism does not take seriously the distinction between persons.

11. Two Principles of Justice

I shall now state in a provisional form the two principles of justice that I believe would be chosen in the original position.

The first statement of the two principles reads as follows.

First: each person is to have an equal right to the most extensive basic liberty compatible with a similar liberty for others.

Second: social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone’s advantage, and (b) attached to positions and offices open to all.

By way of general comment, these principles primarily apply, as I have said, to the basic structure of society. They are to govern the assignment of rights and duties and to regulate the distribution of social and economic advantages. As their formulation suggests, these principles presuppose that the social structure can be divided into two or less distinct parts, the first principle applying to the one, the second to the other. They distinguish between those aspects of the social system that define and secure the equal liberties of citizenship and those that specify and establish social and economic inequalities. The basic liberties of citizens are, roughly speaking, political liberty (the right to vote and to be eligible for public office) together with freedom of speech and assembly; liberty of conscience and freedom of thought; freedom of the person along with the right to hold (personal) property; and freedom from arbitrary arrest and seizure as defined by the concept of the rule of law. These liberties are all required to be equal by the first principle, since citizens of a just society are to have the same basic rights.

The second principle applies, in the first approximation, to the distribution of income and wealth and to the design of organizations that make use of differences in authority and responsibility, or chains of command. While the distribution of wealth and income need not be equal, it must be to everyone’s advantage, and at the same time, positions of authority and offices of command must be accessible to all. One applies the second principle by holding positions open, and then, subject to this constraint, arranges social and economic inequalities so that everyone benefits.

These principles are to be arranged in a serial order with the first principle prior to the second. This ordering means that a departure from the institutions of equal liberty required by the first principle cannot be justified by, or compensated for, by greater social and economic advantages. The distribution of wealth and income, and the hierarchies of authority, must be consistent with both the liberties of equal citizenship and equality of opportunity.

It is clear that these principles are rather specific in their content, and their acceptance rests on certain assumptions that I must eventually try to explain and justify. A theory of justice depends upon a theory of society in ways that will become evident as we proceed. For the present, it should be observed that the two principles (and this holds for all formulations) are a special case of a more general conception of justice that can be expressed as follows.

All social values—liberty and opportunity, income and wealth, and the bases of self-respect—are to be distributed equally unless an unequal distribution of any, or all, of these values is to everyone’s advantage.

Injustice, then, is simply inequalities that are not to the benefit of all. Of course, this conception is extremely vague and requires interpretation.

As a first step, suppose that the basic structure of society distributes certain primary goods, that is, things that every rational man is presumed to want. These goods normally have a use whatever a person’s rational plan of life. For simplicity, assume that the chief primary goods at the disposition of society are rights and liberties, powers and opportunities, income and wealth. (Later on in Part Three the primary good of self-respect has a central place.) These are the social primary goods. Other primary goods such as health and vigor, intelligence and imagination, are natural goods; although their possession is influenced by the basic structure, they are not so directly under its control. Imagine, then, a hypothetical initial arrangement in which all the social primary goods are equally distributed: everyone has similar rights and duties, and income and wealth are evenly shared. This state of affairs provides a benchmark for judging improvements. If certain inequalities of wealth and organizational powers would make everyone better off than in this hypothetical starting situation, then they accord with the general conception.

Now it is possible, at least theoretically, that by giving up some of their fundamental liberties men are sufficiently compensated by the resulting social and economic gains. The general conception of justice imposes no restrictions on what sort of inequalities are permissible; it only requires that everyone’s position be improved. We need not suppose anything so drastic as consenting to a condition of slavery. Imagine instead that men forego certain political rights which the economic returns are significant and their capacity to influence the course of policy by the exercise of these rights would be marginal in any case. It is this kind of exchange which the two principles as stated rule out; being arranged in serial order they do not permit exchanges between basic liberties and economic and social gains. The serial ordering of principles expresses an underlying preference among primary social goods. When this preference is rational so likewise is the choice of these principles in this order.

Now the second principle insists that each person benefit from permissible inequalities in the basic structure. This means that it must be reasonable for each representative man defined by this structure, when he views it as a going concern, to prefer his prospects with the inequality to his prospects without it. One is not allowed to justify differences in income or organizational powers on the ground that the disadvantages of those in one position are outweighed by the greater advantages of those in another. Much less can infringements of liberty be counterbalanced in this way. Applied to the basic structure, the principle of utility would have us maximize the sum of expectations of representative men (weighted by the number of persons they represent, on the classical view); and this would permit us to compensate for the losses of some by the gains of others. Instead, the two principles require that everyone benefit from economic and social inequalities.

26. The Reasoning Leading to the Two Principles of Justice

In this section I take up the choice between the two principles of justice and the principle of average utility. Determining the rational preference between these two options is perhaps the central problem in developing the conception of justice as fairness as a viable alternative to the utilitarian tradition. I shall
begin in this section by presenting some intuitive remarks favoring the two principles. I shall also discuss briefly the qualitative structure of the argument that needs to be made if the case for these principles is to be conclusive.

It will be recalled that the general conception of justice as fairness requires that all primary social goods be distributed equally unless an unequal distribution would be to everyone's advantage. No restrictions are placed on exchanges of these goods and therefore a lesser liberty can be compensated for by greater social and economic benefits. Now looking at the situation from the standpoint of one person selected arbitrarily, there is no way for him to win special advantages for himself. No, on the other hand, are there grounds for his acquiescing in special disadvantages. Since it is not reasonable for him to expect more than an equal share in the division of social goods, and since it is not rational for him to agree to less, the sensible thing for him to do is to acknowledge as the first principle of justice one requiring an equal distribution. Indeed, this principle is so obvious that we would expect it to occur to anyone immediately.

Thus, the parties start with a principle establishing equal liberty for all, including equality of opportunity, as well as an equal distribution of income and wealth. But there is no reason why this acknowledgment should be final. 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. If, for example, these inequalities set up various incentives which succeed in eliciting more productive efforts, a person in the original position may look upon them as necessary to cover the costs of training and to encourage effective performance. One might think that ideally individuals should want to serve one another. But since the parties are assumed not to take an interest in one another's interests, their acceptance of these inequalities is only the acceptance of the relations in which men stand in the circumstances of justice. They have no grounds for complaining of one another's motives. A person in the original position would, therefore, concede the justice of these inequalities. Indeed, it would be shortsighted of him not to do so. He would hesitate to agree to these regularities only if he would be denied by the bare knowledge or perception that others were better situated; and I have assumed that the parties decide as if they are not moved by envy. In order to make the principle regulating inequalities determinate, one looks at the system from the standpoint of the least advantaged representative man. Inequalities are permissible when they maximize, or at least all contribute to, the long-term expectations of the least fortunate group in society.

Now this general conception imposes no constraints on what sorts of inequalities are allowed, whereas the special conception, by putting the two principles in serial order (with the necessary adjustments in meaning) forbids exchanges between basic liberties and economic and social benefits. I shall not try to justify this ordering here. But roughly, the idea underlying this ordering is that if the parties assume that their basic liberties can be effectively exercised, they will not exchange a lesser liberty for an improvement in economic well-being. It is only when social conditions do not allow the effective establishment of these rights that one can conceal their limitation; and these restrictions can be granted only to the extent that they are necessary to prepare the way for a free society. The denial of equal liberty can be defended only if it is necessary to raise the level of civilization so that in due course these freedoms can be enjoyed. Thus, in adopting a serial order we are in effect making a special assumption in the original position, namely, that the parties know that the conditions of their society, whatever they are, are the effective realization of the equal liberties. The serial ordering of the two principles of justice eventually comes to be reasonable if the general conception is consistently followed. This lexical ranking is the long-run tendency of the general view. For the most part I shall assume that the requisite circumstances for the serial order obtain.

It seems clear from these remarks that the two principles are at least a plausible conception of justice. The question, though, is how one is to argue for them more systematically. Now there are several things to do. One can work out their consequences for institutions and note their implications for fundamental social policy. In this way they are tested by a comparison with our considered judgments of justice. But one can also try to find arguments in their favor that are decisive from the standpoint of the original position. In order to see how this might be done, it is useful as a heuristic device to think of the two principles as the maximin solution to the problem of social justice. There is an analogy between the two principles and the maximin rule for choice under uncertainty. This is evident from the fact that the two principles are those a person would choose for the design of a society in which he is to assign his place. The maximin rule tells us to rank alternatives by their worst possible outcomes: we are to adopt the alternative the worst outcome of which is superior to the worst outcome of the others. The persons in the original position do not, of course, assume that their initial place in society is decided by a malevolent opponent. As I note below, they should not reason from false premises. The veil of ignorance does not violate this idea, since an absence of information is not misinformation. But that the two principles of justice would be chosen if the parties were forced to protect themselves against such a contingency explains the sense in which this conception is the maximin solution. And this analogy suggests that if the original position has been described so that it is rational for the parties to adopt the conservative attitude expressed by this rule, a conclusive argument can indeed be constructed for these principles. Clearly the maximin rule is not, in general, a suitable guide for choices under uncertainty. But it is attractive in situations marked by certain special features. My aim, then, is to show that a good case can be made for the two principles based on the fact that the original position manifests these features to the fullest possible degree, carrying them to the limit, so to speak.

Consider the gain-and-loss table below. It represents the gains and losses for a situation which is not a game of strategy. There is no one playing against the person making the decision; instead he is faced with several possible circumstances which may or may not obtain. Which circumstances happen to exist does not depend upon what the person choosing decides or whether he announces his moves in advance. The numbers in the table are monetary values (in hundreds of dollars) in comparison with some initial situation. The gain (g) depends upon the individual's decision (d) and the circumstances (c). Thus g = f(d, c). Assuming that there are three possible decisions and three possible circumstances, we might have this gain-and-loss table.

<table>
<thead>
<tr>
<th>Decision</th>
<th>Circumstance 1</th>
<th>Circumstance 2</th>
<th>Circumstance 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>d_1</td>
<td>-7</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>d_2</td>
<td>-8</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>d_3</td>
<td>5</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

The maximin rule requires that we make the third decision. For in this case the worst that can happen is that one gains five hundred dollars, which is better than the worst for the other actions. If we adopt one of these we may lose either eight or seven hundred dollars. Thus, the choice of d_3 maximizes f(c), for that value of f(c), which for a given d_i, minimizes f. The term "maximin" means the maximum minimum; and the rule directs our attention to the worst that can happen under any proposed course of action, and to decide in the light of that.

Now there appear to be three chief features of situations that give plausibility to this unusual rule. 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. Offhand, the most natural rule of choice would seem to be to compute the expectation of monetary gain for each decision and then to adopt the course of action with the highest prospect. (This expectation is defined as follows: let us suppose that g_j represent the numbers in the gain-and-loss table, where i is the row index and j is the column index; and let pi, j = 1, 2, 3, be the likelihoods of the circumstances, with Σpi = 1. Then the expectation for the i-th decision is equal to Σpg_i.) Thus it must be, for example, that the situation is one in which a knowledge of likelihood is impossible, or at best extremely insecure. In this case it is unreasonable not to be skeptical of probabilistic calculations unless there is no other way out, particularly if the decision
is a fundamental one that needs to be justified to others.

The second feature that suggests the maximin rule is the following: the person choosing has a conception of the good such that he cares very little, if anything, for what he might gain above the maximum stipend that he can, in fact, be sure of by following the maximin rule. It is not worthwhile for him to take a chance for the sake of a further advantage, especially when it may turn out that he loses much that is important to him. This last provision brings in the third feature, namely, that the rejected alternatives have outcomes that one can hardly accept. The situation involves grave risks. Of course these features work most effectively in combination. The paradigm situation for following the maximin rule is when all three features are realized to the highest degree. This rule does not, then, generally apply, nor of course is it self-evident. Rather, it is a maxim, a rule of thumb, that comes into its own in special circumstances. Its application depends upon the qualitative structure of the possible gains and losses in relation to one's conception of the good, all this against a background in which it is reasonable to discount conjectural estimates of likelihoods.

It should be noted, as the comments on the gain-and-loss table say, that the entries in the table represent monetary values and not utilities. This difference is significant since for one thing computing expectations on the basis of such objective values is not the same thing as computing expected utility and may lead to different results. The essential point though is that in justice as fairness the parties do not know their conception of the good and cannot estimate their utility in the ordinary sense. In any case, we want to go behind de facto preferences generated by given conditions. Therefore expectations are based upon an index of primary goods and the parties make their choice accordingly. The entries in the example are in terms of money and not utility to indicate this aspect of the contract doctrine.

Now, as I have suggested, the original position has been defined so that it is a situation in which the maximin rule applies. In order to see this, let us review briefly the nature of this situation with these three special features in mind. To begin with, the veil of ignorance excludes all but the vaguest knowledge of likelihoods. The parties have no basis for determining the probable nature of their society, or their place in it. Thus they have strong reasons for being wary of probability calculations if any other course is open to them. They must also take into account the fact that their choice of principles should seem reasonable to others, in particular their descendants, whose rights will be deeply affected by it. There are further grounds for discounting that I shall mention as we go along. For the present it suffices to note that these considerations are strengthened by the fact that the parties know very little about the gain-and-loss table. Not only are they unable to conjecture the likelihoods of the various possible circumstances, they cannot say much about what the possible circumstances are, much less enumerate them and foresee the outcome of each alternative available. Those deciding are much more in the dark than the illustration by a numerical table suggests. It is for this reason that I have spoken of an analogy with the maximin rule.

Several kinds of arguments for the two principles of justice illustrate the second feature. Thus, if we can maintain that these principles provide a workable theory of social justice, and that they are compatible with reasonable demands of efficiency, then this conception guarantees a satisfactory minimum. There may be, on reflection, little reason for trying to do better.

Finally, the third feature holds if we can assume that other conceptions of justice may lead to institutions that the parties would find intolerable. For example, it has sometimes been held that under some conditions the utility principle (in either form) justifies, if not slavery or scroful, at any rate serious infractions of liberty for the sake of greater social benefits. We need not consider here the truth of this claim, or the likelihood that the requisite conditions obtain. For the moment, this contention is only to illustrate the way in which conceptions of justice may allow for outcomes which the parties may not be able to accept. And having the ready alternative of the two principles of justice which secure a satisfactory minimum, it seems to me if not irrational, for them to take a chance that these outcomes are not realized.

---

**Justice and Entitlement**

**ROBERT NOZICK**

Chapter 7

**Distributive Justice**

The term "distribution of justice" is not a neutral one. Hearing the term "distribution," most people presume that some thing or mechanism uses some principle or criterion to give out a supply of things. Into this process of distributing shares some error may have crept. So it is an open question, at least, whether redistribution should take place; whether we should do again what has already been done once, though poorly. However, we are not in the position of children who have been given portions of pie by someone who now makes last minute adjustments to rectify careless cutting. There is no central distribution, no person or group entitled to control all the resources.


---

**STUDY QUESTIONS**

1. How does Rawls compare justice in social institutions to truth in systems of thought? Do you think this comparison is apt? Why or why not?
2. Explain, in your own words, Rawls's idea of "justice as fairness."
3. What role does the idea of the "veil of ignorance" play in Rawls's theory?
4. What's the idea behind "reflective equilibrium"? Do you think it is a worthy idea? Why or why not?
5. Rawls puts forth two principles of justice that he thinks "would be chosen in the original position." Do you agree with him that these principles would be chosen? Can you think of other principles that might fare better?