

---

A THEORY OF JUSTICE

JOHN RAWLS

---

THE BELKNAP PRESS OF  
HARVARD UNIVERSITY PRESS  
CAMBRIDGE, MASSACHUSETTS

---

For Mard

© Copyright 1971 by the President and Fellows of Harvard College  
All rights reserved  
20 19 18 17 16  
Library of Congress Catalog Card Number 73-168432  
ISBN 0-674-88014-5  
Printed in the United States of America

tion of social cooperation from which it derives. But in doing this we should not lose sight of the special role of the principles of justice of the primary subject to which they apply.

In these preliminary remarks I have distinguished the concept of justice as meaning a proper balance between competing claims from a conception of justice as a set of related principles for identifying the relevant considerations which determine this balance. I have also characterized justice as but one part of a social ideal, although the theory I shall propose no doubt extends its everyday sense. This theory is not offered as a description of ordinary meanings but as an account of certain distributive principles for the basic structure of society. I assume that any reasonably complete ethical theory must include principles for this fundamental problem and that these principles, whatever they are, constitute its doctrine of justice. The concept of justice I take to be defined, then, by the role of 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.

Now this approach may not seem to tally with tradition. I believe, though, that it does. The more specific sense that Aristotle gives to justice, and from which the most familiar formulations derive, is that of refraining from *pleonexia*, that is, from gaining some advantage for oneself by seizing what belongs to another, his property, his reward, his office, and the like, or by denying a person that which is due to him, the fulfillment of a promise, the repayment of a debt, the showing of proper respect, and so on.<sup>3</sup> It is evident that this definition is framed to apply to actions, and persons are thought to be just insofar as they have, as one of the permanent elements of their character, a steady and effective desire to act justly. Aristotle's definition clearly presupposes, however, an account of what properly belongs to a person and of what is due to him. Now such entitlements are, I believe, very often derived from social institutions and the legitimate expectations to which they give rise. There is no reason to

3. *Nicomachean Ethics*, 1129b-1130b5. I have followed the interpretation of Gregory Vlastos, "Justice and Happiness in *The Republic*," in *Plato: A Collection of Critical Essays*, edited by Vlastos (Garden City, N.Y., Doubleday and Company, 1971), vol. 2, pp. 70f. For a discussion of Aristotle on justice, see W. F. R. Hardie, *Aristotle's Ethical Theory* (Oxford, The Clarendon Press, 1968), ch. X.

think that Aristotle would disagree with this, and certainly he has a conception of social justice to account for these claims. The definition I adopt is designed to apply directly to the most important case, the justice of the basic structure. There is no conflict with the traditional notion.

### 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.<sup>4</sup> 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. 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

4. As the text suggests, I shall regard Locke's *Second Treatise of Government*, Rousseau's *The Social Contract*, and Kant's ethical works beginning with *The Foundations of the Metaphysics of Morals* as definitive of the contract tradition. For all of its greatness, Hobbes's *Leviathan* raises special problems. A general historical survey is provided by J. W. Gough, *The Social Contract*, 2nd ed. (Oxford, The Clarendon Press, 1957), and Otto Gierke, *Natural Law and the Theory of Society*, trans. with an introduction by Ernest Barker (Cambridge, The University Press, 1934). A presentation of the contract view as primarily an ethical theory is to be found in G. R. Grice, *The Grounds of Moral Judgment* (Cambridge, The University Press, 1967). See also §19, note 30.

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 culture. It is understood as a purely hypothetical situation characterized so as to lead to a certain conception of justice.<sup>5</sup> 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 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. 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 bargain. 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 agreements reached in it are 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 con-~~

5. Kant is clear that the original agreement is hypothetical. See *The Metaphysics of Morals*, pt. I (*Rechtslehre*), especially §§47, 52; and pt. II of the essay "Concerning the Common Saying: This May Be True in Theory but It Does Not Apply in Practice," in *Kant's Political Writings*, ed. Hans Reiss and trans. by H. B. Nisbet (Cambridge, The University Press, 1970), pp. 73-87. See Georges Vlachos, *La Pensée politique de Kant* (Paris, Presses Universitaires de France, 1962), pp. 326-335; and J. G. Murphy, *Kant: The Philosophy of Right* (London, Macmillan, 1970), pp. 109-112, 133-136, for a further discussion.

~~cepts 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, all 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 would 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.

gether. In particular, I shall try to show further on (§ 17) that this principle is not subject to the objection that it leads to a meritocratic society. Here I wish to consider a few other points, especially its relation to the idea of pure procedural justice.

First, though, I should note that the reasons for requiring open positions are not solely, or even primarily, those of efficiency. I have not maintained that offices must be open if in fact everyone is to benefit from an arrangement. For it may be possible to improve everyone's situation by assigning certain powers and benefits to positions despite the fact that certain groups are excluded from them. Although access is restricted, perhaps these offices can still attract superior talent and encourage better performance. But the principle of open positions forbids this. It expresses the conviction that if some places were not open on a basis fair to all, those kept out would be right in feeling unjustly treated even though they benefited from the greater efforts of those who were allowed to hold them. They would be justified in their complaint not only because they were excluded from certain external rewards of office such as wealth and privilege, but because they were debarred from experiencing the realization of self which comes from a skillful and devoted exercise of social duties. They would be deprived of one of the main forms of human good.

Now I have said that the basic structure is the primary subject of justice. This means, as we have seen, that the first distributive problem is the assignment of fundamental rights and duties and the regulation of social and economic inequalities and of the legitimate expectations founded on these. Of course, any ethical theory recognizes the importance of the basic structure as a subject of justice, but not all theories regard its importance in the same way. In justice as fairness society is interpreted as a cooperative venture for mutual advantage. The basic structure is a public system of rules defining a scheme of activities that leads men to act together so as to produce a greater sum of benefits and assigns to each certain recognized claims to a share in the proceeds. What a person does depends upon what the public rules say he will be entitled to, and what a person is entitled to depends on what he does. The distribution which results is arrived at by honoring the claims determined by what persons undertake to do in the light of these legitimate expectations.

These considerations suggest the idea of treating the question of

distributive shares as a matter of pure procedural justice.<sup>14</sup> The intuitive idea is to design the social system so that the outcome is just whatever it happens to be, at least so long as it is within a certain range. The notion of pure procedural justice is best understood by a comparison with perfect and imperfect procedural justice. To illustrate the former, consider the simplest case of fair division. A number of men are to divide a cake: assuming that the fair division is an equal one, which procedure, if any, will give this outcome? Technicalities aside, the obvious solution is to have one man divide the cake and get the last piece, the others being allowed their pick before him. He will divide the cake equally, since in this way he assures for himself the largest share possible. This example illustrates the two characteristic features of perfect procedural justice. First, there is an independent criterion for what is a fair division, a criterion defined separately from and prior to the procedure which is to be followed. And second, it is possible to devise a procedure that is sure to give the desired outcome. Of course, certain assumptions are made here, such as that the man selected can divide the cake equally, wants as large a piece as he can get, and so on. But we can ignore these details. The essential thing is that there is an independent standard for deciding which outcome is just and a procedure guaranteed to lead to it. Pretty clearly, perfect procedural justice is rare, if not impossible, in cases of much practical interest.

Imperfect procedural justice is exemplified by a criminal trial. The desired outcome is that the defendant should be declared guilty if and only if he has committed the offense with which he is charged. The trial procedure is framed to search for and to establish the truth in this regard. But it seems impossible to design the legal rules so that they always lead to the correct result. The theory of trials examines which procedures and rules of evidence, and the like, are best calculated to advance this purpose consistent with the other ends of the law. Different arrangements for hearing cases may reasonably be expected in different circumstances to yield the right results, not

14. For a general discussion of procedural justice, see Brian Barry, *Political Argument* (London, Routledge and Kegan Paul, 1965), ch. VI. On the problem of fair division, see R. D. Luce and Howard Raiffa, *Games and Decisions* (New York, John Wiley and Sons, Inc., 1957), pp. 363-368; and Hugo Steinhaus, "The Problem of Fair Division," *Econometrica*, vol. 16 (1948).

always but at least most of the time. A trial, then, is an instance of imperfect procedural justice. Even though the law is carefully followed, and the proceedings fairly and properly conducted, it may reach the wrong outcome. An innocent man may be found guilty, a guilty man may be set free. In such cases we speak of a miscarriage of justice: the injustice springs from no human fault but from a fortuitous combination of circumstances which defeats the purpose of the legal rules. The characteristic mark of imperfect procedural justice is that while there is an independent criterion for the correct outcome, there is no feasible procedure which is sure to lead to it.

By contrast, 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. This situation is illustrated by gambling. If a number of persons engage in a series of fair bets, the distribution of cash after the last bet is fair, or at least not unfair, whatever this distribution is. I assume here that fair bets are those having a zero expectation of gain, that the bets are made voluntarily, that no one cheats, and so on. The betting procedure is fair and freely entered into under conditions that are fair. Thus the background circumstances define a fair procedure. Now any distribution of cash summing to the initial stock held by all individuals could result from a series of fair bets. In this sense all of these particular distributions are equally fair. A distinctive feature of pure procedural justice is that the procedure for determining the just result must actually be carried out; for in these cases there is no independent criterion by reference to which a definite outcome can be known to be just. Clearly we cannot say that a particular state of affairs is just because it could have been reached by following a fair procedure. This would permit far too much and would lead to absurdly unjust consequences. It would allow one to say that almost any distribution of goods is just, or fair, since it could have come about as a result of fair gambles. What makes the final outcome of betting fair, or not unfair, is that it is the one which has arisen after a series of fair gambles. A fair procedure translates its fairness to the outcome only when it is actually carried out.

In order, therefore, to apply the notion of pure procedural justice to distributive shares it is necessary to set up and to administer im-

partially a just system of institutions. Only against the background of a just basic structure, including a just political constitution and a just arrangement of economic and social institutions, can one say that the requisite just procedure exists. In Part Two I shall describe in some detail a basic structure that has the necessary features. Its various institutions are explained and connected with the two principles of justice. The intuitive idea is familiar. Suppose that law and government act effectively to keep markets competitive, resources fully employed, property and wealth (especially if private ownership of the means of production is allowed) widely distributed by the appropriate forms of taxation, or whatever, and to guarantee a reasonable social minimum. Assume also that there is fair equality of opportunity underwritten by education for all; and that the other equal liberties are secured. Then it would appear that the resulting distribution of income and the pattern of expectations will tend to satisfy the difference principle. In this complex of institutions, which we think of as establishing social justice in the modern state, the advantages of the better situated improve the condition of the least favored. Or when they do not, they can be adjusted to do so, for example, by setting the social minimum at the appropriate level. As these institutions presently exist they are riddled with grave injustices. But there presumably are ways of running them compatible with their basic design and intention so that the difference principle is satisfied consistent with the demands of liberty and fair equality of opportunity. It is this fact which underlies our assurance that these arrangements can be made just.

It is evident that the role of the principle of fair opportunity is to insure that the system of cooperation is one of pure procedural justice. Unless it is satisfied, distributive justice could not be left to take care of itself, even within a restricted range. Now the great practical advantage of pure procedural justice is that it is no longer necessary in meeting the demands of justice to keep track of the endless variety of circumstances and the changing relative positions of particular persons. One avoids the problem of defining principles to cope with the enormous complexities which would arise if such details were relevant. It is a mistake to focus attention on the varying relative positions of individuals and to require that every change, considered as a single transaction viewed in isolation, be in itself

just. It is the arrangement of the basic structure which is to be judged, and judged from a general point of view. Unless we are prepared to criticize it from the standpoint of a relevant representative man in some particular position, we have no complaint against it. Thus the acceptance of the two principles constitutes an understanding to discard as irrelevant as a matter of social justice much of the information and many of the complications of everyday life.

In pure procedural justice, then, distributions of advantages are not appraised in the first instance by confronting a stock of benefits available with given desires and needs of known individuals. The allotment of the items produced takes place in accordance with the public system of rules, and this system determines what is produced, how much is produced, and by what means. It also determines legitimate claims the honoring of which yields the resulting distribution. Thus in this kind of procedural justice the correctness of the distribution is founded on the justice of the scheme of cooperation from which it arises and on answering the claims of individuals engaged in it. A distribution cannot be judged in isolation from the system of which it is the outcome or from what individuals have done in good faith in the light of established expectations. If it is asked in the abstract whether one distribution of a given stock of things to definite individuals with known desires and preferences is better than another, then there is simply no answer to this question. The conception of the two principles does not interpret the primary problem of distributive justice as one of allocative justice.

By contrast the allocative conception of justice seems naturally to apply when a given collection of goods is to be divided among definite individuals with known desires and needs. The goods to be allotted are not produced by these individuals, nor do these individuals stand in any existing cooperative relations. Since there are no prior claims on the things to be distributed, it is natural to share them out according to desires and needs, or even to maximize the net balance of satisfaction. Justice becomes a kind of efficiency, unless equality is preferred. Suitably generalized, the allocative conception leads to the classical utilitarian view. For as we have seen, this doctrine assimilates justice to the benevolence of the impartial spectator and the latter in turn to the most efficient design of institutions

to promote the greatest balance of satisfaction. As I observed earlier, on this conception society is thought of as so many separate individuals each defining a separate line along which rights and duties are to be assigned and scarce means of satisfaction allocated in accordance with rules so as to give the most complete fulfillment of desire. I shall put aside consideration of the other aspects of this notion until later. The point to note here is that utilitarianism does not interpret the basic structure as a scheme of pure procedural justice. For the utilitarian has, in principle anyway, an independent standard for judging all distributions, namely, whether they produce the greatest net balance of satisfaction. In his theory, institutions are more or less imperfect arrangements for bringing about this end. Thus given existing desires and preferences, and the natural continuations into the future which they allow, the statesman's aim is to set up those social schemes that will best approximate an already specified goal. Since these arrangements are subject to the unavoidable constraints and hindrances of everyday life, the basic structure is a case of imperfect procedural justice.

For the time being I shall suppose that the two parts of the second principle are lexically ordered. Thus we have one lexical ordering within another. But when necessary, this ordering can be modified in the light of the general conception of justice. The advantage of the special conception is that it has a definite shape and suggests certain questions for investigation, for example, under what conditions if any would the lexical ordering be chosen? Our inquiry is given a particular direction and is no longer confined to generalities. Of course, this conception of distributive shares is obviously a great simplification. It is designed to characterize in a clear way a basic structure that makes use of the idea of pure procedural justice. But all the same we should attempt to find simple concepts that can be assembled to give a reasonable conception of justice. The notions of the basic structure, of the veil of ignorance, of a lexical order, of the least favored position, as well as of pure procedural justice are all examples of this. By themselves none of these could be expected to work, but properly put together they may serve well enough. It is too much to suppose that there exists for all or even most moral problems a reasonable solution. Perhaps only a few can be satisfac-

right form when they exemplify the general conception of justice as it is to be interpreted in the light of the difference principle and the lexical ordering to which it tends. Infringements of fair equality of opportunity are not justified by a greater sum of advantages enjoyed by others or by society as a whole. The claim (whether correct or not) must be that the opportunities of the least favored sectors of the community would be still more limited if these inequalities were removed. One is to hold that they are not unjust, since the conditions for achieving the full realization of the principles of justice do not exist.

~~Having noted these cases of priority, I now wish to give the final statement of the two principles of justice for institutions. For the sake of completeness, I shall give a full statement including earlier formulations.~~

#### *First Principle*

Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all.

#### *Second Principle*

Social and economic inequalities are to be arranged so that they are both:

- (a) to the greatest benefit of the least advantaged, consistent with the just savings principle, and
- (b) attached to offices and positions open to all under conditions of fair equality of opportunity.

#### *First Priority Rule (The Priority of Liberty)*

The principles of justice are to be ranked in lexical order and therefore liberty can be restricted only for the sake of liberty. There are two cases:

- (a) a less extensive liberty must strengthen the total system of liberty shared by all;
- (b) a less than equal liberty must be acceptable to those with the lesser liberty.

#### *Second Priority Rule (The Priority of Justice over Efficiency and Welfare)*

The second principle of justice is lexically prior to the principle of efficiency and to that of maximizing the sum of advantages; and

fair opportunity is prior to the difference principle. There are two cases:

- (a) an inequality of opportunity must enhance the opportunities of those with the lesser opportunity;
- (b) an excessive rate of saving must on balance mitigate the burden of those bearing this hardship.

#### *General Conception*

All social primary goods—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 goods is to the advantage of the least favored.

~~By way of comment, these principles and priority rules are no doubt incomplete. Other modifications will surely have to be made, but I shall not further complicate the statement of the principles. It suffices to observe that when we come to nonideal theory, we do not fall back straightway upon the general conception of justice. The lexical ordering of the two principles, and the valuations that this ordering implies, suggest priority rules which seem to be reasonable enough in many cases. By various examples I have tried to illustrate how these rules can be used and to indicate their plausibility. Thus the ranking of the principles of justice in ideal theory reflects back and guides the application of these principles to nonideal situations. It identifies which limitations need to be dealt with first. The drawback of the general conception of justice is that it lacks the definite structure of the two principles in serial order. In more extreme and tangled instances of nonideal theory there may be no alternative to it. At some point the priority of rules for nonideal cases will fail; and indeed, we may be able to find no satisfactory answer at all. But we must try to postpone the day of reckoning as long as possible, and try to arrange society so that it never comes.~~

## 47. THE PRECEPTS OF JUSTICE

~~The sketch of the system of institutions that satisfies the two principles of justice is now complete. Once the just rate of savings is~~



first-person dictatorship and the free-rider forms, since in each case a proper name, or pronoun, or a rigged definite description is needed, either to single out the dictator or to characterize the free-rider. Generality does not, however, exclude general egoism, for each person is allowed to do whatever, in his judgment, is most likely to further his own aims. The principle here can clearly be expressed in a perfectly general way. It is the ordering condition which renders general egoism inadmissible, for if everyone is authorized to advance his aims as he pleases, or if everyone ought to advance his own interests, competing claims are not ranked at all and the outcome is determined by force and cunning.

The several kinds of egoism, then, do not appear on the list presented to the parties. They are eliminated by the formal constraints. Of course, this is not a surprising conclusion, since it is obvious that by choosing one of the other conceptions the persons in the original position can do much better for themselves. Once they ask which principles all should agree to, no form of egoism is a serious candidate for consideration in any case. This only confirms what we knew already, namely, that although egoism is logically consistent and in this sense not irrational, it is incompatible with what we intuitively regard as the moral point of view. The significance of egoism philosophically is not as an alternative conception of right but as a challenge to any such conception. In justice as fairness this is reflected in the fact that we can interpret general egoism as the no-agreement point. It is what the parties would be stuck with if they were unable to reach an understanding.

## 24. THE VEIL OF IGNORANCE

The idea of the original position 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. Somehow we must nullify the effects of specific contingencies which put men at odds and tempt them to exploit social and natural circumstances to their own advantage. Now in order to do this I assume that the parties are situated behind a veil of ignorance. They do not know how the various alternatives will affect their own particular case

and they are obliged to evaluate principles solely on the basis of general considerations.<sup>11</sup>

It is assumed, then, that the parties do not know certain kinds of particular facts. First of all, no one knows his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength, and the like. Nor, again, does anyone know his conception of the good, the particulars of his rational plan of life, or even the special features of his psychology such as his aversion to risk or liability to optimism or pessimism. More than this, I assume that the parties do not know the particular circumstances of their own society. That is, they do not know its economic or political situation, or the level of civilization and culture it has been able to achieve. The persons in the original position have no information as to which generation they belong. These broader restrictions on knowledge are appropriate in part because questions of social justice arise between generations as well as within them, for example, the question of the appropriate rate of capital saving and of the conservation of natural resources and the environment of nature. There is also, theoretically anyway, the question of a reasonable genetic policy. In these cases too, in order to carry through the idea of the original position, the parties must not know the contingencies that set them in opposition. They must choose principles the consequences of which they are prepared to live with whatever generation they turn out to belong to.

As far as possible, then, the only particular facts which the parties know is that their society is subject to the circumstances of justice and whatever this implies. It is taken for granted, however, that they know the general facts about human society. They understand political affairs and the principles of economic theory; they know the basis of social organization and the laws of human psychology. Indeed, the parties are presumed to know whatever general facts affect the choice of the principles of justice. There are

11. The veil of ignorance is so natural a condition that something like it must have occurred to many. The closest explicit statement of it known to me is found in J. C. Harsanyi, "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking," *Journal of Political Economy*, vol. 61 (1953). Harsanyi uses it to develop a utilitarian theory, as I discuss below in §§27-28.

no limitations on general information, that is, on general laws and theories, since conceptions of justice must be adjusted to the characteristics of the systems of social cooperation which they are to regulate, and there is no reason to rule out these facts. It is, for example, a consideration against a conception of justice that, in view of the laws of moral psychology, men would not acquire a desire to act upon it even when the institutions of their society satisfied it. For in this case there would be difficulty in securing the stability of social cooperation. It is an important feature of a conception of justice that it should generate its own support. That is, its principles should be such that when they are embodied in the basic structure of society men tend to acquire the corresponding sense of justice. Given the principles of moral learning, men develop a desire to act in accordance with its principles. In this case a conception of justice is stable. This kind of general information is admissible in the original position.

The notion of the veil of ignorance raises several difficulties. Some may object that the exclusion of nearly all particular information makes it difficult to grasp what is meant by the original position. Thus it may be helpful to observe that one or more persons can at any time enter this position, or perhaps, better, simulate the deliberations of this hypothetical situation simply by reasoning in accordance with the appropriate restrictions. In arguing for a conception of justice we must be sure that it is among the permitted alternatives and satisfies the stipulated formal constraints. No considerations can be advanced in its favor unless they would be rational ones for us to urge were we to lack the kind of knowledge that is excluded. The evaluation of principles must proceed in terms of the general consequences of their public recognition and universal application, it being assumed that they will be complied with by everyone. To say that a certain conception of justice would be chosen in the original position is equivalent to saying that rational deliberation satisfying certain conditions and restrictions would reach a certain conclusion. If necessary, the argument to this result could be set out more formally. I shall, however, speak throughout in terms of the notion of the original position. It is more economical and suggestive, and brings out certain essential features that otherwise one might easily overlook.

These remarks show that the original position is not to be thought of as a general assembly which includes at one moment everyone who will live at some time; or, much less, as an assembly of everyone who could live at some time. It is not a gathering of all actual or possible persons. To conceive of the original position in either of these ways is to stretch fantasy too far; the conception would cease to be a natural guide to intuition. In any case, it is important that the original position be interpreted so that one can at any time adopt its perspective. It must make no difference when one takes up this viewpoint, or who does so; the restrictions must be such that the same principles are always chosen. The veil of ignorance is a key condition in meeting this requirement. It insures not only that the information available is relevant, but that it is at all times the same.

It may be protested that the condition of the veil of ignorance is irrational. Surely, some may object, principles should be chosen in the light of all the knowledge available. There are various replies to this contention. Here I shall sketch those which emphasize the simplifications that need to be made if one is to have any theory at all. (Those based on the Kantian interpretation of the original position are given later, § 40.) To begin with, it is clear that since the differences among the parties are unknown to them, and everyone is equally rational and similarly situated, each is convinced by the same arguments. Therefore, we can view the choice in the original position from the standpoint of one person selected at random. If anyone after due reflection prefers a conception of justice to another, then they all do, and a unanimous agreement can be reached. We can, to make the circumstances more vivid, imagine that the parties are required to communicate with each other through a referee as intermediary, and that he is to announce which alternatives have been suggested and the reasons offered in their support. He forbids the attempt to form coalitions, and he informs the parties when they have come to an understanding. But such a referee is actually superfluous, assuming that the deliberations of the parties must be similar.

Thus there follows the very important consequence that the parties have no basis for bargaining in the usual sense. No one knows his situation in society nor his natural assets, and therefore no one is in a position to tailor principles to his advantage. We

might imagine that one of the contractees threatens to hold out unless the others agree to principles favorable to him. But how does he know which principles are especially in his interests? The same holds for the formation of coalitions: if a group were to decide to band together to the disadvantage of the others, they would not know how to favor themselves in the choice of principles. Even if they could get everyone to agree to their proposal, they would have no assurance that it was to their advantage, since they cannot identify themselves either by name or description. The one case where this conclusion fails is that of saving. Since the persons in the original position know that they are contemporaries (taking the present time of entry interpretation), they can favor their generation by refusing to make any sacrifices at all for their successors; they simply acknowledge the principle that no one has a duty to save for posterity. Previous generations have saved or they have not; there is nothing the parties can now do to affect that. So in this instance the veil of ignorance fails to secure the desired result. Therefore I resolve the question of justice between generations in a different way by altering the motivation assumption. But with this adjustment no one is able to formulate principles especially designed to advance his own cause. Whatever his temporal position, each is forced to choose for everyone.<sup>12</sup>

The restrictions on particular information in the original position are, then, of fundamental importance. Without them we would not be able to work out any definite theory of justice at all. We would have to be content with a vague formula stating that justice is what would be agreed to without being able to say much, if anything, about the substance of the agreement itself. The formal constraints of the concept of right, those applying to principles directly, are not sufficient for our purpose. The veil of ignorance makes possible a unanimous choice of a particular conception of justice. Without these limitations on knowledge the bargaining problem of the original position would be hopelessly complicated. Even if theoretically a solution were to exist, we would not, at present anyway, be able to determine it.

The notion of the veil of ignorance is implicit, I think, in Kant's

12. Rousseau, *The Social Contract*, bk. II, ch. IV, par. 5.

ethics (§40). Nevertheless the problem of defining the knowledge of the parties and of characterizing the alternatives open to them has often been passed over, even by contract theories. Sometimes the situation definitive of moral deliberation is presented in such an indeterminate way that one cannot ascertain how it will turn out. Thus Perry's doctrine is essentially contractarian: he holds that social and personal integration must proceed by entirely different principles, the latter by rational prudence, the former by the concurrence of persons of good will. He would appear to reject utilitarianism on much the same grounds suggested earlier: namely, that it improperly extends the principle of choice for one person to choices facing society. The right course of action is characterized as that which best advances social aims as these would be formulated by reflective agreement given that the parties have full knowledge of the circumstances and are moved by a benevolent concern for one another's interests. No effort is made, however, to specify in any precise way the possible outcomes of this sort of agreement. Indeed, without a far more elaborate account, no conclusions can be drawn.<sup>13</sup> I do not wish here to criticize others; rather, I want to explain the necessity for what may seem at times like so many irrelevant details.

Now the reasons for the veil of ignorance go beyond mere simplicity. We want to define the original position so that we get the desired solution. If a knowledge of particulars is allowed, then the outcome is biased by arbitrary contingencies. As already observed, to each according to his threat advantage is not a principle of justice. If the original position is to yield agreements that are just, the parties must be fairly situated and treated equally as moral persons. The arbitrariness of the world must be corrected for by adjusting the circumstances of the initial contractual situation. Moreover, if in choosing principles we required unanimity even when there is full information, only a few rather obvious cases could be decided. A conception of justice based on unanimity in these circumstances would indeed be weak and trivial. But once knowledge is excluded, the requirement of unanimity is not out of place and the fact that

13. See R. B. Perry, *The General Theory of Value* (New York, Longmans, Green and Company, 1926), pp. 674-682.

it can be satisfied is of great importance. It enables us to say of the preferred conception of justice that it represents a genuine reconciliation of interests.

A final comment. For the most part I shall suppose that the parties possess all general information. No general facts are closed to them. I do this mainly to avoid complications. Nevertheless a conception of justice is to be the public basis of the terms of social cooperation. Since common understanding necessitates certain bounds on the complexity of principles, there may likewise be limits on the use of theoretical knowledge in the original position. Now clearly it would be very difficult to classify and to grade for complexity the various sorts of general facts. I shall make no attempt to do this. We do however recognize an intricate theoretical construction when we meet one. Thus it seems reasonable to say that other things equal one conception of justice is to be preferred to another when it is founded upon markedly simpler general facts, and its choice does not depend upon elaborate calculations in the light of a vast array of theoretically defined possibilities. It is desirable that the grounds for a public conception of justice should be evident to everyone when circumstances permit. This consideration favors, I believe, the two principles of justice over the criterion of utility.

## 25. THE RATIONALITY OF THE PARTIES

I have assumed throughout that the persons in the original position are rational. In choosing between principles each tries as best he can to advance his interests. But I have also assumed that the parties do not know their conception of the good. This means that while they know that they have some rational plan of life, they do not know the details of this plan, the particular ends and interests which it is calculated to promote. How, then, can they decide which conceptions of justice are most to their advantage? Or must we suppose that they are reduced to mere guessing? To meet this difficulty, I postulate that they accept the account of the good touched upon in the preceding chapter: they assume that they would prefer more primary social goods rather than less. Of course, it may turn out once the veil of ignorance is removed, that some of them for religious or other reasons may not, in fact, want more of these goods.

But from the standpoint of the original position, it is rational for the parties to suppose that they do want a larger share, since in any case they are not compelled to accept more if they do not wish to, nor does a person suffer from a greater liberty. Thus even though the parties are deprived of information about their particular ends, they have enough knowledge to rank the alternatives. They know that in general they must try to protect their liberties, widen their opportunities, and enlarge their means for promoting their aims whatever these are. Guided by the theory of the good and the general facts of moral psychology, their deliberations are no longer guesswork. They can make a rational decision in the ordinary sense.

The concept of rationality invoked here, with the exception of one essential feature, is the standard one familiar in social theory.<sup>14</sup> Thus in the usual way, a rational person is thought to have a coherent set of preferences between the options open to him. He ranks these options according to how well they further his purposes; he follows the plan which will satisfy more of his desires rather than less, and which has the greater chance of being successfully executed. The special assumption I make is that a rational individual does not suffer from envy. He is not ready to accept a loss for himself if only others have less as well. He is not downcast by the knowledge or perception that others have a larger index of primary social goods. Or at least this is true as long as the differences between himself and others do not exceed certain limits, and he does not believe that the existing inequalities are founded on injustice or are the result of letting chance work itself out for no compensating social purpose (§ 80).

14. For this notion of rationality, see the references to Sen and Arrow above, § 23, note 9. The discussion in I. M. D. Little, *The Critique of Welfare Economics*, 2nd ed. (Oxford, Clarendon Press, 1957), ch. II, is also relevant here. For rational choice under uncertainty, see below § 26, note 18. H. A. Simon discusses the limitations of the classical conceptions of rationality and the need for a more realistic theory in "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics*, vol. 69 (1955). See also his essay in *Surveys of Economic Theory*, vol. 3 (London, Macmillan, 1967). For philosophical discussions see Donald Davidson, "Actions, Reasons, and Causes," *Journal of Philosophy*, vol. 60 (1963); C. G. Hempel, *Aspects of Scientific Explanation* (New York, The Free Press, 1965), pp. 463-486; Jonathan Bennett, *Rationality* (London, Routledge and Kegan Paul, 1964), and J. D. Mabbott, "Reason and Desire," *Philosophy*, vol. 28 (1953).

The assumption that the parties are not moved by envy raises certain questions. Perhaps we should also assume that they are not liable to various other feelings such as shame and humiliation (§ 67). Now a satisfactory account of justice will eventually have to deal with these matters too, but for the present I shall leave these complications aside. Another objection to our procedure is that it is too unrealistic. Certainly men are afflicted with these feelings. How can a conception of justice ignore this fact? I shall meet this problem by dividing the argument for the principles of justice into two parts. In the first part, the principles are derived on the supposition that envy does not exist; while in the second, we consider whether the conception arrived at is feasible in view of the circumstances of human life.

One reason for this procedure is that envy tends to make everyone worse off. In this sense it is collectively disadvantageous. Presuming its absence amounts to supposing that in the choice of principles men should think of themselves as having their own plan of life which is sufficient for itself. They have a secure sense of their own worth so that they have no desire to abandon any of their aims provided others have less means to further theirs. I shall work out a conception of justice on this stipulation to see what happens. Later I shall try to show that when the principles adopted are put into practice, they lead to social arrangements in which envy and other destructive feelings are not likely to be strong. The conception of justice eliminates the conditions that give rise to disruptive attitudes. It is, therefore, inherently stable (§§ 80–81).

The assumption of mutually disinterested rationality, then, comes to this: the persons in the original position try to acknowledge principles which advance their system of ends as far as possible. They do this by attempting to win for themselves the highest index of primary social goods, since this enables them to promote their conception of the good most effectively whatever it turns out to be. The parties do not seek to confer benefits or to impose injuries on one another; they are not moved by affection or rancor. Nor do they try to gain relative to each other; they are not envious or vain. Put in terms of a game, we might say: they strive for as high an absolute score as possible. They do not wish a high or a low score for their opponents, nor do they seek to maximize or minimize the differ-

ence between their successes and those of others. The idea of a game does not really apply, since the parties are not concerned to win but to get as many points as possible judged by their own system of ends.

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 does not mean that in their deliberations the parties apply some particular conception of justice, for this would defeat the point of the motivation assumption. Rather, 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. It is essential to observe, however, that this assumption still permits the consideration of men's capacity to act on the various conceptions of justice. The general facts of human psychology and the principles of moral learning are relevant matters for the parties to examine. If a conception of justice is unlikely to generate its own support, or lacks stability, this fact must not be overlooked. For then a different conception of justice might be preferred. The assumption only says that the parties have a capacity for justice in a purely formal sense: taking everything relevant into account, including the general facts of moral psychology, the parties will adhere to the principles eventually chosen. They are rational in that they will not enter into agreements they know they cannot keep, or can do so only with great difficulty. Along with other considerations, they count the strains of commitment (§ 29). Thus in assessing conceptions of justice the persons in the original position are to assume that the one they adopt will be strictly complied with. The consequences of their agreement are to be worked out on this basis.

~~With the preceding remarks about rationality and motivation of the parties the description of the original position is for the most part complete. We can summarize this description with the following list of elements of the initial situation and their varia-~~

Principles. (The asterisks mark the interpretations that constitute the original position.)

1. The Nature of the Parties (§ 22)
  - \*a. continuing persons (family heads, or genetic lines)
  - b. single individuals
  - c. associations (states, churches, or other corporate bodies)
2. Subject of Justice (§ 2)
  - \*a. basic structure of society
  - b. rules of corporate associations
  - c. law of nations
3. Presentation of Alternatives (§ 21)
  - \*a. shorter (or longer) list
  - b. general characterization of the possibilities
4. Time of Entry (§ 24)
  - \*a. any time (during age of reason) for living persons
  - b. all actual persons (those alive at some time) simultaneously
  - c. all possible persons simultaneously
5. Circumstances of Justice (§ 22)
  - \*a. Hume's conditions of moderate scarcity
  - b. the above plus further extremes
6. Formal Conditions on Principles (§ 23)
  - \*a. generality, universality, publicity, ordering, and finality
  - b. the above less publicity, say
7. Knowledge and Beliefs (§ 24)
  - \*a. veil of ignorance
  - b. full information
  - c. partial knowledge
8. Motivation of the Parties (§ 25)
  - \*a. mutual disinterestedness (limited altruism)
  - b. elements of social solidarity and good will
  - c. perfect altruism
9. Rationality (§§ 25, 28)
  - \*a. taking effective means to ends with unified expectations and objective interpretation of probability
  - b. as above but without unified expectations and using the principle of insufficient reason

10. Agreement Condition (§ 24)
  - \*a. unanimity in perpetuity
  - b. majority acceptance, or whatever, for limited period
11. Compliance Condition (§ 25)
  - \*a. strict compliance
  - b. partial compliance in various degrees
12. No Agreement Point (§ 23)
  - \*a. general egoism
  - b. the state of nature

We can turn now to the choice of principles. But first I shall mention a few misunderstandings to be avoided. First of all, we must keep in mind that the parties in the original position are theoretically defined individuals. The grounds for their consent are set out by the description of the contractual situation and their preference for primary goods. Thus to say that the principles of justice would be adopted is to say how these persons would decide being moved in ways our account describes. Of course, when we try to simulate the original position in everyday life, that is, when we try to conduct ourselves in moral argument as its constraints require, we will presumably find that our deliberations and judgments are influenced by our special inclinations and attitudes. Surely it will prove difficult to correct for our various propensities and aversions in striving to adhere to the conditions of this idealized situation. But none of this affects the contention that in the original position rational persons so characterized would make a certain decision. This proposition belongs to the theory of justice. It is another question how well human beings can assume this role in regulating their practical reasoning.

Since the persons in the original position are assumed to take no interest in one another's interests (although they may have a concern for third parties), it may be thought that justice as fairness is itself an egoistic theory. It is not, of course, one of the three forms of egoism mentioned earlier, but some may think, as Schopenhauer thought of Kant's doctrine, that it is egoistic nevertheless.<sup>15</sup> Now this is a misconception. For the fact that in the original

15. See *On the Basis of Ethics* (1840), trans. E. F. J. Payne (New York, The Liberal Arts Press, Inc., 1965), pp. 89-92.

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. Clearly the two principles of justice and the principles of obligation and natural duty require us to consider the rights and claims of others. And the sense of justice is a normally effective desire to comply with these restrictions. The motivation of the persons in the original position must not be confused with the motivation of persons in everyday life who accept the principles that would be chosen and who have the corresponding sense of justice. In practical affairs an individual does have a knowledge of his situation and he can, if he wishes, exploit contingencies to his advantage. Should his sense of justice move him to act on the principles of right that would be adopted in the original position, his desires and aims are surely not egoistic. He voluntarily takes on the limitations expressed by this interpretation of the moral point of view.

This conclusion is supported by a further reflection. Once we consider the idea of a contract theory 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, or an interest in one another's interests. Perry, as I mentioned before, thinks of the right standards and decisions as those promoting the ends reached by reflective agreement under circumstances making for impartiality and good will. Now the combination of mutual disinterest and the veil of ignorance achieves the same purpose as benevolence. For this combination of conditions forces each person in the original position to take the good of others into account. In justice as fairness, then, the effects of good will are brought about by several conditions working jointly. The feeling that this conception of justice is egoistic is an illusion fostered by looking at but one of the elements of the original position. Furthermore, this pair of assumptions has enormous advantages over that of benevolence plus knowledge. As I have noted, the latter is so complex that no definite theory at all can be worked out. Not only are the complications caused by so much information insurmountable, but the motivational assumption requires clarification. For example, what is the relative strength of benevolent desires? In brief, the combination of mutual disin-

terestedness plus the veil of ignorance has the merits of simplicity and clarity while at the same time insuring the effects of what are at first sight morally more attractive assumptions. And if it is asked why one should not postulate benevolence with the veil of ignorance, the answer is that there is no need for so strong a condition. Moreover, it would defeat the purpose of grounding the theory of justice on weak stipulations, as well as being incongruous with the circumstances of justice.

Finally, if the parties are conceived as themselves making proposals, they have no incentive to suggest pointless or arbitrary principles. For example, none would urge that special privileges be given to those exactly six feet tall or born on a sunny day. Nor would anyone put forward the principle that basic rights should depend on the color of one's skin or the texture of one's hair. No one can tell whether such principles would be to his advantage. Furthermore, each such principle is a limitation of one's liberty of action, and such restrictions are not to be accepted without a reason. Certainly we might imagine peculiar circumstances in which these characteristics are relevant. Those born on a sunny day might be blessed with a happy temperament, and for some positions of authority this might be a qualifying attribute. But such distinctions would never be proposed in first principles, for these must have some rational connection with the advancement of human interests broadly defined. The rationality of the parties and their situation in the original position guarantees that ethical principles and conceptions of justice have this general content.<sup>16</sup> Inevitably, then, racial and sexual discrimination presupposes that some hold a favored place in the social system which they are willing to exploit to their advantage. From the standpoint of persons similarly situated in an initial situation which is fair, the principles of explicit racist doctrines are not only unjust. They are irrational. For this reason we could say that they are not moral conceptions at all, but simply means of suppression. They have no place on a reason-

16. For a different way of reaching this conclusion, see Philippa Foot, "Moral Arguments," *Mind*, vol. 67 (1958), and "Moral Beliefs," *Proceedings of the Aristotelian Society*, vol. 59 (1958-1959); and R. W. Beardsmore, *Moral Reasoning* (New York, Schocken Books, 1969), especially ch. IV. The problem of content is discussed briefly in G. F. Warnock, *Contemporary Moral Philosophy* (London, Macmillan, 1967), pp. 55-61.

able list of traditional conceptions of justice.<sup>17</sup> Of course, this contention is not at all a matter of definition. It is rather a consequence of the conditions characterizing the original position, especially the conditions of the rationality of the parties and the veil of ignorance. That conceptions of right have a certain content and exclude arbitrary and pointless principles is, therefore, an inference from the theory.

## 26. THE REASONING LEADING TO THE TWO PRINCIPLES OF JUSTICE

In this and the next two sections 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. Nor, 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.

17. For a similar view, see B. A. O. Williams, "The Idea of Equality," *Philosophy, Politics, and Society*, Second Series, ed. Peter Laslett and W. G. Runciman (Oxford, Basil Blackwell, 1962), p. 113.

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 dejected 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. From time to time in later chapters this problem will be considered (§§ 39, 82). 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 concede 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, admit 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. Part II is devoted to this. 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.<sup>18</sup> This is evident from the fact that the two principles are those a person would choose for the design of a society in which his enemy is to assign him his place. The maximin rule tells us to rank alternatives by their worst possible

18. An accessible discussion of this and other rules of choice under uncertainty can be found in W. J. Baumol, *Economic Theory and Operations Analysis*, 2nd ed. (Englewood Cliffs, N.J., Prentice-Hall Inc., 1965), ch. 24. Baumol gives a geometric interpretation of these rules, including the diagram used in §13 to illustrate the difference principle. See pp. 558-562. See also R. D. Luce and Howard Raiffa, *Games and Decisions* (New York, John Wiley and Sons, Inc., 1957), ch. XIII, for a fuller account.

possible outcomes: we are to adopt the alternative the worst outcome of which is superior to the worst outcomes 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.

Decisions	Circumstances		
	$c_1$	$c_2$	$c_3$
$d_1$	-7	8	12
$d_2$	-8	7	14
$d_3$	5	6	8

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  $f(d,c)$  for that value of  $c$ , which for a given  $d$ , minimizes  $f$ . The term "maximin" means the *maximum minimorum*; 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.<sup>19</sup> 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  $g_{ij}$  represent the numbers in the gain-and-loss table, where  $i$  is the row index and  $j$  is the column index; and let  $p_j$ ,  $j = 1, 2, 3$ , be the likelihoods of the circumstances, with  $\sum p_j = 1$ . Then the expectation for the  $i$ th decision is equal to  $\sum p_j g_{ij}$ .) Thus it must be, for example, that the situation is one in which a knowledge of likelihoods 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 minimum 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 combina-

tion. 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 enu-

19. Here I borrow from William Fellner, *Probability and Profit* (Homewood, Ill., R. D. Irwin, Inc., 1965), pp. 140-142, where these features are noted.