A Theory of Justice

JOHN RAWLS

John Rawls is James B. Conant University Professor, Emeritus, at Harvard University. After publishing A Theory of Justice in 1971 he continued to rethink his views on justice, resulting in his later books Political Liberalism and The Law of Peoples.

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 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 culture. 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 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 jus-
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.”

**FURTHER CASES OF PRIORITY**

I now wish to give the final statement of the two principles of justice for institutions.

*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 the basic liberties can be restricted only for the sake of liberty. There are two cases:

(a) a less extensive liberty must strengthen the total system of liberties 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.

**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.

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 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. An important feature of a conception of justice is that it should generate its own support. 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 and 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 favor of a conception of justice 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. If we conceived of the original position in either of these ways, the conception would cease to be a natural guide to intuition and would lack a clear sense. In any case, the original position must 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. . . . 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 agreement 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, to handle the question of justice between generations, I modify the motivation assumption and add a further constraint. With these adjustments, no generation is able to formulate principles especially designed to advance its own cause and some significant limits on savings principles can be derived. Whatever a person’s temporal position, each is forced to choose for all.4

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 ethics. 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 contractual: 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. 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 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 the complexity of 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.

THE RATIONALITY OF THE PARTIES

I have assumed throughout that the persons in the original position are rational. But I have also assumed that they 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 normally 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. 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. 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.

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. 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.

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 difference 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. 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.

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 forms of although they may have a concern for third parties), first agreement arc to be worked out on this basis.

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 of justice 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. Thus, more generally, the motivation of the parties in the original position does not determine directly the motivation of people in a just society. For in the latter case, we assume that its members grow up and live under a just basic structure, as the two principles require; and then we try to work out what kind of conception of the good and moral sentiments people would acquire. Therefore the mutual disinterestedness of the parties determines other motivations only indirectly, that is, via its effects on the agreement on principles. It is these principles, together with the laws of psychology (as these work under the conditions of just institutions), which shape the aims and moral sentiments of citizens of a well-ordered society.

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 much 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 disinterestedness 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.

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. 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 reasonable list of traditional conceptions of justice. 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.

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.

Now consider the point of view of anyone in the original position. 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 primary goods, and since it is not rational for him to agree to less, the sensible thing is to acknowledge as the first step a principle of justice requiring an equal distribution. Indeed, this principle is so obvious given the symmetry of the parties that it would occur to everyone immediately. Thus the parties start with a principle requiring equal basic liberties for all, as well as fair equality of opportunity and equal division of income and wealth.

But even holding firm to the priority of the basic liberties and fair equality of opportunity, there is no reason why this initial acknowledgment should be final. Society should take into account economic efficiency and the requirements of organization and technology. If there are inequalities in income and wealth, and differences in authority and degrees of responsibility, that work to make everyone better off in comparison with the benchmark of equality, why
not permit them? One might think that ideally individuals should want to serve one another. But since the parties are assumed to be mutually disinterested, their acceptance of these economic and institutional inequalities is only the recognition of the relations of opposition in which men stand in the circumstances of justice. They have no grounds for complaining of one another’s motives. Thus the parties would agree to these differences only if they would be dejected by the bare knowledge or perception that others are better situated; but I suppose that they decide as if they are not moved by envy. Thus the basic structure should allow these inequalities so long as these improve everyone’s situation, including that of the least advantaged, provided that they are consistent with equal liberty and fair opportunity. Because the parties start from an equal division of all social primary goods, those who benefit least have, so to speak, a veto. Thus we arrive at the difference principle. Taking equality as the basis of comparison, those who have gained more must do so on terms that are justifiable to those who have gained the least.

By some such reasoning, then, the parties might arrive at the two principles of justice in serial order. I shall not try to justify this ordering here, but the following remarks may convey the intuitive idea. I assume that the parties view themselves as free persons who have fundamental aims and interests in the name of which they think it legitimate for them to make claims on one another concerning the design of the basic structure of society. The religious interest is a familiar historical example; the interest in the integrity of the person is another. In the original position the parties do not know what particular forms these interests take; but they do assume that they have such interests and that the basic liberties necessary for their protection are guaranteed by the first principle. Since they must secure these interests, they rank the first principle prior to the second. The case for the two principles can be strengthened by spelling out in more detail the notion of a free person. Very roughly the parties regard themselves as having a highest-order interest in how all their other interests, including even their fundamental ones, are shaped and regulated by social institutions. They do not think of themselves as inevitably bound to, or as identical with, the pursuit of any particular complex of fundamental interests that they may have at any given time, although they want the right to advance such interests (provided they are admissible). Rather, free persons conceive of themselves as beings who can revise and alter their final ends and who give first priority to preserving their liberty in these matters. Hence, they not only have final ends that they are in principle free to pursue or to reject, but their original allegiance and continued devotion to these ends are to be formed and affirmed under conditions that are free. Since the two principles secure a social form that maintains these conditions, they would be agreed to rather than the principle of utility. Only by this agreement can the parties be sure that their highest-order interest as free persons is guaranteed.

The priority of liberty means that whenever the basic liberties can be effectively established, a lesser or an unequal liberty cannot be exchanged for an improvement in economic well-being. It is only when social circumstances do not allow the effective establishment of these basic rights that one can concede their limitation; and even then these restrictions can be granted only to the extent that they are necessary to prepare the way for the time when they are no longer justified. The denial of the equal liberties can be defended only when it is essential to change the conditions of civilization so that in due course these liberties can be enjoyed. Thus in adopting the serial order of the two principles, the parties are assuming that the conditions of their society, whatever they are, admit the effective realization of the equal liberties. Or that if they do not, circumstances are nevertheless sufficiently favorable so that the priority of the first principle points out the most urgent changes and identifies the preferred path to the social state in which all the basic liberties can be fully instituted. The complete realization of the two principles in serial order is the long-run tendency of this ordering, at least under reasonably fortunate conditions.

It seems 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 a relation 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 his enemy is to assign him 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 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 maximum rule is not, in general, a suitable guide for choices under uncertainty. But it holds only 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 has these features to a very high degree.

NOTES


3. The veil of ignorance is so natural a condition that something like it must have occurred to many. The formulation in the text is implicit. I believe, in Kant's doctrine of the categorical imperative, both in the way this procedural criterion is defined and the use Kant makes of it. Thus when Kant tells us to test our maxim by considering what would be the case were it a universal law of nature, he must suppose that we do not know our place within this imagined system of nature. See, for example, his discussion of the topic of practical judgment in The Critique of Practical Reason, Academy Edition, vol. 5, pp. 68–72. A similar restriction on information 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). However, other aspects of Harsanyi's view are quite different, and he uses the restriction to develop a utilitarian theory. . . .


11. 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>Circumstances</th>
<th>c_1</th>
<th>c_2</th>
<th>c_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 (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.