
A THEORY OF JUSTICE

Revised Edition

JOHN RAWLS

THE BELKNAP PRESS OF
HARVARD UNIVERSITY PRESS
CAMBRIDGE, MASSACHUSETTS

© Copyright 1971, 1999 by the President and Fellows of Harvard College

All rights reserved

Printed in the United States of America

This book is a revised edition of *A Theory of Justice*,
published in 1971 by Harvard University Press.

Library of Congress Cataloging-in-Publication Data

Rawls, John, 1921–

A theory of justice / John Rawls. — Rev. ed.

p. cm.

Includes bibliographical references and index.

ISBN 0-674-00077-3 (cloth : alk. paper). —

ISBN 0-674-00078-1 (paper : alk. paper)

1. Justice. I. Title.

JC578.R38 1999

320'.01'1—dc21 99-29110

CHAPTER I. JUSTICE AS FAIRNESS

In this introductory chapter I sketch some of the main ideas of the theory of justice I wish to develop. The exposition is informal and intended to prepare the way for the more detailed arguments that follow. Unavoidably there is some overlap between this and later discussions. I begin by describing the role of justice in social cooperation and with a brief account of the primary subject of justice, the basic structure of society. I then present the main idea of justice as fairness, a theory of justice that generalizes and carries to a higher level of abstraction the traditional conception of the social contract. The compact of society is replaced by an initial situation that incorporates certain procedural constraints on arguments designed to lead to an original agreement on principles of justice. I also take up, for purposes of clarification and contrast, the classical utilitarian and intuitionist conceptions of justice and consider some of the differences between these views and justice as fairness. My guiding aim is to work out a theory of justice that is a viable alternative to these doctrines which have long dominated our philosophical tradition.

1. THE ROLE OF JUSTICE

Justice is the first virtue of social institutions, as truth is of systems of thought. A theory however elegant and economical must be rejected or revised if it is untrue; likewise laws and institutions no matter how efficient and well-arranged must be reformed or abolished if they are unjust. Each person possesses an inviolability founded on justice that even the welfare of society as a whole cannot override. For this reason justice denies that the loss of freedom for some is made right by a greater good shared by others. It does not allow that the sacrifices imposed on a few are outweighed by the larger sum of advantages enjoyed by many. Therefore in a just society the liberties of equal citizenship are taken as settled;

the rights secured by justice are not subject to political bargaining or to the calculus of social interests. The only thing that permits us to acquiesce in an erroneous theory is the lack of a better one; analogously, an injustice is tolerable only when it is necessary to avoid an even greater injustice. Being first virtues of human activities, truth and justice are uncompromising.

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

Now let us say that a society is well-ordered when it is not only designed to advance the good of its members but when it is also effectively regulated by a public conception of justice. That is, it is a society in which (1) everyone accepts and knows that the others accept the same principles of justice, and (2) the basic social institutions generally satisfy and are generally known to satisfy these principles. In this case while men may put forth excessive demands on one another, they nevertheless acknowledge a common point of view from which their claims may be adjudicated. If men's inclination to self-interest makes their vigilance

against one another necessary, their public sense of justice makes their secure association together possible. Among individuals with disparate aims and purposes a shared conception of justice establishes the bonds of civic friendship; the general desire for justice limits the pursuit of other ends. One may think of a public conception of justice as constituting the fundamental charter of a well-ordered human association.

Existing societies are of course seldom well-ordered in this sense, for what is just and unjust is usually in dispute. Men disagree about which principles should define the basic terms of their association. Yet we may still say, despite this disagreement, that they each have a conception of justice. That is, they understand the need for, and they are prepared to affirm, a characteristic set of principles for assigning basic rights and duties and for determining what they take to be the proper distribution of the benefits and burdens of social cooperation. Thus it seems natural to think of the concept of justice as distinct from the various conceptions of justice and as being specified by the role which these different sets of principles, these different conceptions, have in common.¹ Those who hold different conceptions of justice can, then, still agree that institutions are just when no arbitrary distinctions are made between persons in the assigning of basic rights and duties and when the rules determine a proper balance between competing claims to the advantages of social life. Men can agree to this description of just institutions since the notions of an arbitrary distinction and of a proper balance, which are included in the concept of justice, are left open for each to interpret according to the principles of justice that he accepts. These principles single out which similarities and differences among persons are relevant in determining rights and duties and they specify which division of advantages is appropriate. Clearly this distinction between the concept and the various conceptions of justice settles no important questions. It simply helps to identify the role of the principles of social justice.

Some measure of agreement in conceptions of justice is, however, not the only prerequisite for a viable human community. There are other fundamental social problems, in particular those of coordination, efficiency, and stability. Thus the plans of individuals need to be fitted together so that their activities are compatible with one another and they can all be carried through without anyone's legitimate expectations being severely disappointed. Moreover, the execution of these plans should lead to the

1. Here I follow H. L. A. Hart, *The Concept of Law* (Oxford, The Clarendon Press, 1961), pp. 155–159.

achievement of social ends in ways that are efficient and consistent with justice. And finally, the scheme of social cooperation must be stable: it must be more or less regularly complied with and its basic rules willingly acted upon; and when infractions occur, stabilizing forces should exist that prevent further violations and tend to restore the arrangement. Now it is evident that these three problems are connected with that of justice. In the absence of a certain measure of agreement on what is just and unjust, it is clearly more difficult for individuals to coordinate their plans efficiently in order to insure that mutually beneficial arrangements are maintained. Distrust and resentment corrode the ties of civility, and suspicion and hostility tempt men to act in ways they would otherwise avoid. So while the distinctive role of conceptions of justice is to specify basic rights and duties and to determine the appropriate distributive shares, the way in which a conception does this is bound to affect the problems of efficiency, coordination, and stability. We cannot, in general, assess a conception of justice by its distributive role alone, however useful this role may be in identifying the concept of justice. We must take into account its wider connections; for even though justice has a certain priority, being the most important virtue of institutions, it is still true that, other things equal, one conception of justice is preferable to another when its broader consequences are more desirable.

2. THE SUBJECT OF JUSTICE

Many different kinds of things are said to be just and unjust: not only laws, institutions, and social systems, but also particular actions of many kinds, including decisions, judgments, and imputations. We also call the attitudes and dispositions of persons, and persons themselves, just and unjust. Our topic, however, is that of social justice. For us the primary subject of justice is the basic structure of society, or more exactly, the way in which the major social institutions distribute fundamental rights and duties and determine the division of advantages from social cooperation. By major institutions I understand the political constitution and the principal economic and social arrangements. Thus the legal protection of freedom of thought and liberty of conscience, competitive markets, private property in the means of production, and the monogamous family are examples of major social institutions. Taken together as one scheme, the major institutions define men's rights and duties and influence their life prospects, what they can expect to be and how well they can hope to

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 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.⁴ 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,

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.

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.⁵ 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 concepts of justice and fairness are the same, any more than the phrase "poetry as metaphor" means that the concepts of poetry and metaphor are the same.

Justice as fairness begins, as I have said, with one of the most general of all choices which persons might make together, namely, with the

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.

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. They are to presume that even their spiritual aims may be opposed, in the way that the aims of those of different religions may be opposed. Moreover, the concept of rationality must be interpreted as far as possible in the narrow sense, standard in economic theory, of taking the most effective means to given ends. I shall modify this concept to some extent, as explained later (§25), but one must try to avoid introducing into it any controversial ethical elements. The initial situation must be characterized by stipulations that are widely accepted.

In working out the conception of justice as fairness one main task

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

I shall maintain instead that the persons in the initial situation would choose two rather different principles: the first requires equality in the assignment of basic rights and duties, while the second holds that social and economic inequalities, for example inequalities of wealth and authority, are just only if they result in compensating benefits for everyone, and in particular for the least advantaged members of society. These principles rule out justifying institutions on the grounds that the hardships of some are offset by a greater good in the aggregate. It may be expedient but it is not just that some should have less in order that others may prosper. But there is no injustice in the greater benefits earned by a few provided that the situation of persons not so fortunate is thereby improved. The intuitive idea is that since everyone's well-being depends upon a scheme of cooperation without which no one could have a satisfactory life, the division of advantages should be such as to draw forth the willing cooperation of everyone taking part in it, including those less well situated. The two principles mentioned seem to be a fair basis on which those better endowed, or more fortunate in their social position, neither of which we can be said to deserve, could expect the willing cooperation of others when some workable scheme is a necessary condition of the wel-

fare of all.⁶ Once we decide to look for a conception of justice that prevents the use of the accidents of natural endowment and the contingencies of social circumstance as counters in a quest for political and economic advantage, we are led to these principles. They express the result of leaving aside those aspects of the social world that seem arbitrary from a moral point of view.

The problem of the choice of principles, however, is extremely difficult. I do not expect the answer I shall suggest to be convincing to everyone. It is, therefore, worth noting from the outset that justice as fairness, like other contract views, consists of two parts: (1) an interpretation of the initial situation and of the problem of choice posed there, and (2) a set of principles which, it is argued, would be agreed to. One may accept the first part of the theory (or some variant thereof), but not the other, and conversely. The concept of the initial contractual situation may seem reasonable although the particular principles proposed are rejected. To be sure, I want to maintain that the most appropriate conception of this situation does lead to principles of justice contrary to utilitarianism and perfectionism, and therefore that the contract doctrine provides an alternative to these views. Still, one may dispute this contention even though one grants that the contractarian method is a useful way of studying ethical theories and of setting forth their underlying assumptions.

Justice as fairness is an example of what I have called a contract theory. Now there may be an objection to the term “contract” and related expressions, but I think it will serve reasonably well. Many words have misleading connotations which at first are likely to confuse. The terms “utility” and “utilitarianism” are surely no exception. They too have unfortunate suggestions which hostile critics have been willing to exploit; yet they are clear enough for those prepared to study utilitarian doctrine. The same should be true of the term “contract” applied to moral theories. As I have mentioned, to understand it one has to keep in mind that it implies a certain level of abstraction. In particular, the content of the relevant agreement is not to enter a given society or to adopt a given form of government, but to accept certain moral principles. Moreover, the undertakings referred to are purely hypothetical: a contract view holds that certain principles would be accepted in a well-defined initial situation.

The merit of the contract terminology is that it conveys the idea that principles of justice may be conceived as principles that would be chosen by rational persons, and that in this way conceptions of justice may be

6. For the formulation of this intuitive idea I am indebted to Allan Gibbard.

explained and justified. The theory of justice is a part, perhaps the most significant part, of the theory of rational choice. Furthermore, principles of justice deal with conflicting claims upon the advantages won by social cooperation; they apply to the relations among several persons or groups. The word "contract" suggests this plurality as well as the condition that the appropriate division of advantages must be in accordance with principles acceptable to all parties. The condition of publicity for principles of justice is also connoted by the contract phraseology. Thus, if these principles are the outcome of an agreement, citizens have a knowledge of the principles that others follow. It is characteristic of contract theories to stress the public nature of political principles. Finally there is the long tradition of the contract doctrine. Expressing the tie with this line of thought helps to define ideas and accords with natural piety. There are then several advantages in the use of the term "contract." With due precautions taken, it should not be misleading.

A final remark. Justice as fairness is not a complete contract theory. For it is clear that the contractarian idea can be extended to the choice of more or less an entire ethical system, that is, to a system including principles for all the virtues and not only for justice. Now for the most part I shall consider only principles of justice and others closely related to them; I make no attempt to discuss the virtues in a systematic way. Obviously if justice as fairness succeeds reasonably well, a next step would be to study the more general view suggested by the name "rightness as fairness." But even this wider theory fails to embrace all moral relationships, since it would seem to include only our relations with other persons and to leave out of account how we are to conduct ourselves toward animals and the rest of nature. I do not contend that the contract notion offers a way to approach these questions which are certainly of the first importance; and I shall have to put them aside. We must recognize the limited scope of justice as fairness and of the general type of view that it exemplifies. How far its conclusions must be revised once these other matters are understood cannot be decided in advance.

4. THE ORIGINAL POSITION AND JUSTIFICATION

I have said that the original position is the appropriate initial status quo which insures that the fundamental agreements reached in it are fair. This fact yields the name "justice as fairness." It is clear, then, that I want to say that one conception of justice is more reasonable than another, or

justifiable with respect to it, if rational persons in the initial situation would choose its principles over those of the other for the role of justice. Conceptions of justice are to be ranked by their acceptability to persons so circumstanced. Understood in this way the question of justification is settled by working out a problem of deliberation: we have to ascertain which principles it would be rational to adopt given the contractual situation. This connects the theory of justice with the theory of rational choice.

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

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

One should not be misled, then, by the somewhat unusual conditions which characterize the original position. The idea here is simply to make vivid to ourselves the restrictions that it seems reasonable to impose on arguments for principles of justice, and therefore on these principles themselves. Thus it seems reasonable and generally acceptable that no one should be advantaged or disadvantaged by natural fortune or social circumstances in the choice of principles. It also seems widely agreed that it should be impossible to tailor principles to the circumstances of one's own case. We should insure further that particular inclinations and aspirations, and persons' conceptions of their good do not affect the prin-

ciples adopted. The aim is to rule out those principles that it would be rational to propose for acceptance, however little the chance of success, only if one knew certain things that are irrelevant from the standpoint of justice. For example, if a man knew that he was wealthy, he might find it rational to advance the principle that various taxes for welfare measures be counted unjust; if he knew that he was poor, he would most likely propose the contrary principle. To represent the desired restrictions one imagines a situation in which everyone is deprived of this sort of information. One excludes the knowledge of those contingencies which sets men at odds and allows them to be guided by their prejudices. In this manner the veil of ignorance is arrived at in a natural way. This concept should cause no difficulty if we keep in mind the constraints on arguments that it is meant to express. At any time we can enter the original position, so to speak, simply by following a certain procedure, namely, by arguing for principles of justice in accordance with these restrictions.

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

There is, however, another side to justifying a particular description of the original position. This is to see if the principles which would be chosen match our considered convictions of justice or extend them in an acceptable way. We can note whether applying these principles would lead us to make the same judgments about the basic structure of society which we now make intuitively and in which we have the greatest confidence; or whether, in cases where our present judgments are in doubt and given with hesitation, these principles offer a resolution which we can affirm on reflection. There are questions which we feel sure must be answered in a certain way. For example, we are confident that religious intolerance and racial discrimination are unjust. We think that we have examined these things with care and have reached what we believe is an

impartial judgment not likely to be distorted by an excessive attention to our own interests. These convictions are provisional fixed points which we presume any conception of justice must fit. But we have much less assurance as to what is the correct distribution of wealth and authority. Here we may be looking for a way to remove our doubts. We can check an interpretation of the initial situation, then, by the capacity of its principles to accommodate our firmest convictions and to provide guidance where guidance is needed.

In searching for the most favored description of this situation we work from both ends. We begin by describing it so that it represents generally shared and preferably weak conditions. We then see if these conditions are strong enough to yield a significant set of principles. If not, we look for further premises equally reasonable. But if so, and these principles match our considered convictions of justice, then so far well and good. But presumably there will be discrepancies. In this case we have a choice. We can either modify the account of the initial situation or we can revise our existing judgments, for even the judgments we take provisionally as fixed points are liable to revision. By going back and forth, sometimes altering the conditions of the contractual circumstances, at others withdrawing our judgments and conforming them to principle, I assume that eventually we shall find a description of the initial situation that both expresses reasonable conditions and yields principles which match our considered judgments duly pruned and adjusted. This state of affairs I refer to as reflective equilibrium.⁷ It is an equilibrium because at last our principles and judgments coincide; and it is reflective since we know to what principles our judgments conform and the premises of their derivation. At the moment everything is in order. But this equilibrium is not necessarily stable. It is liable to be upset by further examination of the conditions which should be imposed on the contractual situation and by particular cases which may lead us to revise our judgments. Yet for the time being we have done what we can to render coherent and to justify our convictions of social justice. We have reached a conception of the original position.

I shall not, of course, actually work through this process. Still, we may think of the interpretation of the original position that I shall present as the result of such a hypothetical course of reflection. It represents the

7. The process of mutual adjustment of principles and considered judgments is not peculiar to moral philosophy. See Nelson Goodman, *Fact, Fiction, and Forecast* (Cambridge, Mass., Harvard University Press, 1955), pp. 65–68, for parallel remarks concerning the justification of the principles of deductive and inductive inference.

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

A final comment. We shall want to say that certain principles of justice are justified because they would be agreed to in an initial situation of equality. I have emphasized that this original position is purely hypothetical. It is natural to ask why, if this agreement is never actually entered into, we should take any interest in these principles, moral or otherwise. The answer is that the conditions embodied in the description of the original position are ones that we do in fact accept. Or if we do not, then perhaps we can be persuaded to do so by philosophical reflection. Each aspect of the contractual situation can be given supporting grounds. Thus what we shall do is to collect together into one conception a number of conditions on principles that we are ready upon due consideration to recognize as reasonable. These constraints express what we are prepared to regard as limits on fair terms of social cooperation. One way to look at the idea of the original position, therefore, is to see it as an expository device which sums up the meaning of these conditions and helps us to extract their consequences. On the other hand, this conception is also an intuitive notion that suggests its own elaboration, so that led on by it we are drawn to define more clearly the standpoint from which we can best interpret moral relationships. We need a conception that enables us to envision our objective from afar: the intuitive notion of the original position is to do this for us.⁸

5. CLASSICAL UTILITARIANISM

There are many forms of utilitarianism, and the development of the theory has continued in recent years. I shall not survey these forms here, nor

8. Henri Poincaré remarks: "Il nous faut une faculté qui nous fasse voir le but de loin, et, cette faculté, c'est l'intuition." *La Valeur de la science* (Paris, Flammarion, 1909), p. 27.

ples that hold unconditionally whatever the circumstances. This fact is connected with the Kantian interpretation of justice as fairness. But I leave this matter aside until later (§40).

Finally, an obvious point. An argument for the two principles, or indeed for any conception, is always relative to some list of alternatives. If we change the list, the argument will, in general, have to be different. A similar sort of remark applies to all features of the original position. There are indefinitely many variations of the initial situation and therefore no doubt indefinitely many theorems of moral geometry. Only a few of these are of any philosophical interest, since most variations are irrelevant from a moral point of view. We must try to steer clear of side issues while at the same time not losing sight of the special assumptions of the argument.

22. THE CIRCUMSTANCES OF JUSTICE

The circumstances of justice may be described as the normal conditions under which human cooperation is both possible and necessary.³ Thus, as I noted at the outset, although a society is a cooperative venture for mutual advantage, it is typically marked by a conflict as well as an identity of interests. There is an identity of interests since social cooperation makes possible a better life for all than any would have if each were to try to live solely by his own efforts. There is a conflict of interests since men are not indifferent as to how the greater benefits produced by their collaboration are distributed, for in order to pursue their ends they each prefer a larger to a lesser share. Thus principles are needed for choosing among the various social arrangements which determine this division of advantages and for underwriting an agreement on the proper distributive shares. These requirements define the role of justice. The background conditions that give rise to these necessities are the circumstances of justice.

These conditions may be divided into two kinds. First, there are the objective circumstances which make human cooperation both possible and necessary. Thus, many individuals coexist together at the same time on a definite geographical territory. These individuals are roughly similar

3. My account largely follows that of Hume in *A Treatise of Human Nature*, bk. III, pt. II, sec. ii, and *An Enquiry Concerning the Principles of Morals*, sec. III, pt. I. But see also H. L. A. Hart, *The Concept of Law* (Oxford, The Clarendon Press, 1961), pp. 189–195, and J. R. Lucas, *The Principles of Politics* (Oxford, The Clarendon Press, 1966), pp. 1–10.

in physical and mental powers; or at any rate, their capacities are comparable in that no one among them can dominate the rest. They are vulnerable to attack, and all are subject to having their plans blocked by the united force of others. Finally, there is the condition of moderate scarcity understood to cover a wide range of situations. Natural and other resources are not so abundant that schemes of cooperation become superfluous, nor are conditions so harsh that fruitful ventures must inevitably break down. While mutually advantageous arrangements are feasible, the benefits they yield fall short of the demands men put forward.

The subjective circumstances are the relevant aspects of the subjects of cooperation, that is, of the persons working together. Thus while the parties have roughly similar needs and interests, or needs and interests in various ways complementary, so that mutually advantageous cooperation among them is possible, they nevertheless have their own plans of life. These plans, or conceptions of the good, lead them to have different ends and purposes, and to make conflicting claims on the natural and social resources available. Moreover, although the interests advanced by these plans are not assumed to be interests in the self, they are the interests of a self that regards its conception of the good as worthy of recognition and that advances claims in its behalf as deserving satisfaction. I also suppose that men suffer from various shortcomings of knowledge, thought, and judgment. Their knowledge is necessarily incomplete, their powers of reasoning, memory, and attention are always limited, and their judgment is likely to be distorted by anxiety, bias, and a preoccupation with their own affairs. Some of these defects spring from moral faults, from selfishness and negligence; but to a large degree, they are simply part of men's natural situation. As a consequence individuals not only have different plans of life but there exists a diversity of philosophical and religious belief, and of political and social doctrines.

Now this constellation of conditions I shall refer to as the circumstances of justice. Hume's account of them is especially perspicuous and the preceding summary adds nothing essential to his much fuller discussion. For simplicity I often stress the condition of moderate scarcity (among the objective circumstances), and that of conflict of interests (among the subjective circumstances). Thus, one can say, in brief, that the circumstances of justice obtain whenever persons put forward conflicting claims to the division of social advantages under conditions of moderate scarcity. Unless these circumstances existed there would be no occasion for the virtue of justice, just as in the absence of threats of injury to life and limb there would be no occasion for physical courage.

Several clarifications should be noted. First of all, I shall, of course, assume that the persons in the original position know that these circumstances of justice obtain. This much they take for granted about the conditions of their society. A further assumption is that the parties try to advance their conception of the good as best they can, and that in attempting to do this they are not bound by prior moral ties to each other.

The question arises, however, whether the persons in the original position have obligations and duties to third parties, for example, to their immediate descendants. To say that they do would be one way of handling questions of justice between generations. However, the aim of justice as fairness is to try to derive all duties and obligations of justice from other reasonable conditions. So, if possible, this way out should be avoided. There are several other courses open to us. We can adopt a motivation assumption and think of the parties as representing a continuing line of claims. For example, we can assume that they are heads of families and therefore have a desire to further the well-being of at least their more immediate descendants. Or we can require the parties to agree to principles subject to the constraint that they wish all preceding generations to have followed the very same principles. By an appropriate combination of such stipulations, I believe that the whole chain of generations can be tied together and principles agreed to that suitably take into account the interests of each (§§24, 44). If this is right, we will have succeeded in deriving duties to other generations from reasonable conditions.

It should be noted that I make no restrictive assumptions about the parties' conceptions of the good except that they are rational long-term plans. While these plans determine the aims and interests of a self, the aims and interests are not presumed to be egoistic or selfish. Whether this is the case depends upon the kinds of ends which a person pursues. If wealth, position, and influence, and the accolades of social prestige, are a person's final purposes, then surely his conception of the good is egoistic. His dominant interests are in himself, not merely, as they must always be, interests of a self.⁴ There is no inconsistency, then, in supposing that once the veil of ignorance is removed, the parties find that they have ties of sentiment and affection, and want to advance the interests of others and to see their ends attained. But the postulate of mutual disinterest in the original position is made to insure that the principles of justice do not depend upon strong assumptions. Recall that the original position is meant to incorporate widely shared and yet weak conditions. A conception of

4. On this point see W. T. Stace, *The Concept of Morals* (London, Macmillan, 1937), pp. 221–223.

justice should not presuppose, then, extensive ties of natural sentiment. At the basis of the theory, one tries to assume as little as possible.

Finally, I shall assume that the parties in the original position are mutually disinterested: they are not willing to have their interests sacrificed to the others. The intention is to model men's conduct and motives in cases where questions of justice arise. The spiritual ideals of saints and heroes can be as irreconcilably opposed as any other interests. Conflicts in pursuit of these ideals are the most tragic of all. Thus justice is the virtue of practices where there are competing interests and where persons feel entitled to press their rights on each other. In an association of saints agreeing on a common ideal, if such a community could exist, disputes about justice would not occur. Each would work selflessly for one end as determined by their common religion, and reference to this end (assuming it to be clearly defined) would settle every question of right. But a human society is characterized by the circumstances of justice. The account of these conditions involves no particular theory of human motivation. Rather, its aim is to reflect in the description of the original position the relations of individuals to one another which set the stage for questions of justice.

23. THE FORMAL CONSTRAINTS OF THE CONCEPT OF RIGHT

The situation of the persons in the original position reflects certain constraints. The alternatives open to them and their knowledge of their circumstances are limited in various ways. These restrictions I refer to as the constraints of the concept of right since they hold for the choice of all ethical principles and not only for those of justice. If the parties were to acknowledge principles for the other virtues as well, these constraints would also apply.

I shall consider first the constraints on the alternatives. There are certain formal conditions that it seems reasonable to impose on the conceptions of justice that are to be allowed on the list presented to the parties. I do not claim that these conditions follow from the concept of right, much less from the meaning of morality. I avoid an appeal to the analysis of concepts at crucial points of this kind. There are many constraints that can reasonably be associated with the concept of right, and different selections can be made from these and counted as definitive within a particular theory. The merit of any definition depends upon the

soundness of the theory that results; by itself, a definition cannot settle any fundamental question.⁵

The propriety of these formal conditions is derived from the task of principles of right in adjusting the claims that persons make on their institutions and one another. If the principles of justice are to play their role, that of assigning basic rights and duties and determining the division of advantages, these requirements are natural enough. Each of them is suitably weak and I assume that they are satisfied by the traditional conceptions of justice. These conditions do, however, exclude the various forms of egoism, as I note below, which shows that they are not without moral force. This makes it all the more necessary that the conditions not be justified by definition or the analysis of concepts, but only by the reasonableness of the theory of which they are a part. I arrange them under five familiar headings.

First of all, principles should be general. That is, it must be possible to formulate them without the use of what would be intuitively recognized as proper names, or rigged definite descriptions. Thus the predicates used in their statement should express general properties and relations. Unfortunately deep philosophical difficulties seem to bar the way to a satisfactory account of these matters.⁶ I shall not try to deal with them here. In presenting a theory of justice one is entitled to avoid the problem of defining general properties and relations and to be guided by what seems reasonable. Further, since the parties have no specific information about themselves or their situation, they cannot identify themselves anyway. Even if a person could get others to agree, he does not know how to tailor principles to his advantage. The parties are effectively forced to stick to general principles, understanding the notion here in an intuitive fashion.

The naturalness of this condition lies in part in the fact that first

5. Various interpretations of the concept of morality are discussed by W. K. Frankena, "Recent Conceptions of Morality," in *Morality and the Language of Conduct*, ed. H. N. Castañeda and George Nakhnikian (Detroit, Wayne State University Press, 1965), and "The Concept of Morality," *Journal of Philosophy*, vol. 63 (1966). The first of these essays contains numerous references. The account in the text is perhaps closest to that of Kurt Baier in *The Moral Point of View* (Ithaca, N.Y., Cornell University Press, 1958), ch. VIII. I follow Baier in emphasizing the conditions of publicity (he does not use this term, but it is implied by his stipulation of universal teachability, pp. 195f), ordering, finality, and material content (although on the contract view the last condition follows as a consequence, see §25 and note 16 below). For other discussions, see R. M. Hare, *The Language of Morals* (Oxford, The Clarendon Press, 1952), W. D. Falk, "Morality, Self, and Others," also in *Morality and the Language of Conduct*, and P. F. Strawson, "Social Morality and Individual Ideal," *Philosophy*, vol. 36 (1961).

6. See, for example, W. V. Quine, *Ontological Relativity and Other Essays* (New York, Columbia University Press, 1969), ch. 5 entitled "Natural Kinds."

principles must be capable of serving as a public charter of a well-ordered society in perpetuity. Being unconditional, they always hold (under the circumstances of justice), and the knowledge of them must be open to individuals in any generation. Thus, to understand these principles should not require a knowledge of contingent particulars, and surely not a reference to individuals or associations. Traditionally the most obvious test of this condition is the idea that what is right is that which accords with God's will. But in fact this doctrine is normally supported by an argument from general principles. For example, Locke held that the fundamental principle of morals is the following: if one person is created by another (in the theological sense), then that person has a duty to comply with the precepts set to him by his creator.⁷ This principle is perfectly general and given the nature of the world on Locke's view, it singles out God as the legitimate moral authority. The generality condition is not violated, although it may appear so at first sight.

Next, principles are to be universal in application. They must hold for everyone in virtue of their being moral persons. Thus I assume that each can understand these principles and use them in his deliberations. This imposes an upper bound of sorts on how complex they can be, and on the kinds and number of distinctions they draw. Moreover, a principle is ruled out if it would be self-contradictory, or self-defeating, for everyone to act upon it. Similarly, should a principle be reasonable to follow only when others conform to a different one, it is also inadmissible. Principles are to be chosen in view of the consequences of everyone's complying with them.

As defined, generality and universality are distinct conditions. For example, egoism in the form of first-person dictatorship (Everyone is to serve my—or Pericles'—interests) satisfies universality but not generality. While all could act in accordance with this principle, and the results might in some cases not be at all bad, depending on the interests of the dictator, the personal pronoun (or the name) violates the first condition. Again, general principles may not be universal. They may be framed to hold for a restricted class of individuals, for instance those singled out by special biological or social characteristics, such as hair color or class situation, or whatever. To be sure, in the course of their lives individuals acquire obligations and assume duties that are peculiar to them. Never-

7. See *Essays on the Laws of Nature*, ed. W. von Leyden (Oxford, The Clarendon Press, 1954), the fourth essay, especially pp. 151–157.

theless these various duties and obligations are the consequence of first principles that hold for all as moral persons; the derivation of these requirements has a common basis.

A third condition is that of publicity, which arises naturally from a contractarian standpoint. The parties assume that they are choosing principles for a public conception of justice.⁸ They suppose that everyone will know about these principles all that he would know if their acceptance were the result of an agreement. Thus the general awareness of their universal acceptance should have desirable effects and support the stability of social cooperation. The difference between this condition and that of universality is that the latter leads one to assess principles on the basis of their being intelligently and regularly followed by everyone. But it is possible that all should understand and follow a principle and yet this fact not be widely known or explicitly recognized. The point of the publicity condition is to have the parties evaluate conceptions of justice as publicly acknowledged and fully effective moral constitutions of social life. The publicity condition is clearly implicit in Kant's doctrine of the categorical imperative insofar as it requires us to act in accordance with principles that one would be willing as a rational being to enact as law for a kingdom of ends. He thought of this kingdom as an ethical commonwealth, as it were, which has such moral principles for its public charter.

A further condition is that a conception of right must impose an ordering on conflicting claims. This requirement springs directly from the role of its principles in adjusting competing demands. There is a difficulty, however, in deciding what counts as an ordering. It is clearly desirable that a conception of justice be complete, that is, able to order all the claims that can arise (or that are likely to in practice). And the ordering

8. Publicity is clearly implied in Kant's notion of the moral law, but the only place I know of where he discusses it expressly is in *Perpetual Peace*, appendix II; see *Political Writings*, ed. Hans Reiss and trans. H. B. Nisbet (Cambridge, The University Press, 1970), pp. 125–130. There are of course brief statements elsewhere. For example, in *The Metaphysics of Morals*, pt. I (*Rechtslehre*), §43, he says: "Public Right is the sum total of those laws which require to be made universally public in order to produce a state of right." In "Theory and Practice" he remarks in a footnote: "No right in a state can be tacitly and treacherously included by a secret reservation, and least of all a right which the people claim to be a part of the constitution, for all laws within it must be thought of as arising out of a public will. Thus if a constitution allowed rebellion, it would have to declare this right publicly and make clear how it might be implemented." *Political Writings*, pp. 136, 84n, respectively. I believe Kant intends this condition to apply to a society's conception of justice. See also note 4, §51, below; and Baier, cited in note 5 above. There is a discussion of common knowledge and its relation to agreement in D. K. Lewis, *Convention* (Cambridge, Mass., Harvard University Press, 1969), esp. pp. 52–60, 83–88.

should in general be transitive: if, say, a first arrangement of the basic structure is ranked more just than a second, and the second more just than a third, then the first should be more just than the third. These formal conditions are natural enough, though not always easy to satisfy.⁹ But is trial by combat a form of adjudication? After all, physical conflict and resort to arms result in an ordering; certain claims do win out over others. The main objection to this ordering is not that it may be intransitive. Rather, it is to avoid the appeal to force and cunning that the principles of right and justice are accepted. Thus I assume that to each according to his threat advantage is not a conception of justice. It fails to establish an ordering in the required sense, an ordering based on certain relevant aspects of persons and their situation which are independent from their social position, or their capacity to intimidate and coerce.¹⁰

The fifth and last condition is that of finality. The parties are to assess the system of principles as the final court of appeal in practical reasoning. There are no higher standards to which arguments in support of claims can be addressed; reasoning successfully from these principles is conclusive. If we think in terms of the fully general theory which has principles for all the virtues, then such a theory specifies the totality of relevant considerations and their appropriate weights, and its requirements are decisive. They override the demands of law and custom, and of social rules generally. We are to arrange and respect social institutions as the princi-

9. For a discussion of orderings and preference relations, see A. K. Sen, *Collective Choice and Social Welfare* (San Francisco, Holden-Day Inc., 1970), chs. 1 and 1*; and K. J. Arrow, *Social Choice and Individual Values*, 2nd ed. (New York, John Wiley, 1963), ch. II.

10. *Theory of Games as a Tool for the Moral Philosopher* (Cambridge, The University Press, 1955). On the analysis he presents, it turns out that the fair division of playing time between Matthew and Luke depends on their preferences, and these in turn are connected with the instruments they wish to play. Since Matthew has a threat advantage over Luke, arising from the fact that Matthew, the trumpeter, prefers both of them playing at once to neither of them playing, whereas Luke, the pianist, prefers silence to cacophony, Matthew is allotted twenty-six evenings of play to Luke's seventeen. If the situation were reversed, the threat advantage would be with Luke. See pp. 36f. But we have only to suppose that Matthew is a jazz enthusiast who plays the drums, and Luke a violinist who plays sonatas, in which case it will be fair on this analysis for Matthew to play whenever and as often as he likes, assuming as it is plausible to assume that he does not care whether Luke plays or not. Clearly something has gone wrong. What is lacking is a suitable definition of a status quo that is acceptable from a moral point of view. We cannot take various contingencies as known and individual preferences as given and expect to elucidate the concept of justice (or fairness) by theories of bargaining. The conception of the original position is designed to meet the problem of the appropriate status quo. A similar objection to Braithwaite's analysis is found in J. R. Lucas, "Moralists and Gamesmen," *Philosophy*, vol. 34 (1959), pp. 9f. For another discussion, consult Sen, *Collective Choice and Social Welfare*, pp. 118–123, who argues that the solution of J. F. Nash in "The Bargaining Problem," *Econometrica*, vol. 18 (1950), is similarly defective from an ethical point of view.

ples of right and justice direct. Conclusions from these principles also override considerations of prudence and self-interest. This does not mean that these principles insist upon self-sacrifice; for in drawing up the conception of right the parties take their interests into account as best they can. The claims of personal prudence are already given an appropriate weight within the full system of principles. The complete scheme is final in that when the course of practical reasoning it defines has reached its conclusion, the question is settled. The claims of existing social arrangements and of self-interest have been duly allowed for. We cannot at the end count them a second time because we do not like the result.

Taken together, then, these conditions on conceptions of right come to this: a conception of right is a set of principles, general in form and universal in application, that is to be publicly recognized as a final court of appeal for ordering the conflicting claims of moral persons. Principles of justice are identified by their special role and the subject to which they apply. Now by themselves the five conditions exclude none of the traditional conceptions of justice. It should be noted, however, that they do rule out the listed variants of egoism. The generality condition eliminates both 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.¹¹

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 re-

11. 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. See the last paragraph of §27.

sources 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 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. 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. (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 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 situ-

ation 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 (§22). 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 (§44). Whatever a person's temporal position, each is forced to choose for all.¹²

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 (§40). Nevertheless the problem of defining the knowledge of the parties and of characterizing the alternatives open to them has often been passed

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

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.¹³ 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

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

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.

25. 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