

Justice, Fairness and Rationality*

Joel Feinberg†

A Theory of Justice. By John Rawls. Cambridge: Harvard University Press, 1971. Pp. 607. \$15.00 (clothbound), \$3.95 (paperback).

The publication of John Rawls' *A Theory of Justice* is an event to be celebrated by philosophers. It is not often we get a book like this one, not a mere sketch or collection of essays but a thoroughly elaborated theory, similar in scope and ambition to the great books of Spinoza, Locke, and Kant. Rawls' book is reminiscent of the seventeenth and eighteenth centuries in other respects too. It is unremittingly rationalistic in spirit. "Moral philosophy," he tells us, "[is a] part of the theory of rational choice,"¹ and in the theory of justice "we should strive for a kind of moral geometry with all the rigor which this name connotes."² Indeed, Rawls is reluctant to allow that equally well informed reasonable men can disagree about anything that is important. Rawls' book is also similar to seventeenth and eighteenth century prototypes in substance, for it is a "social contract theory," he tells us, in the tradition of Rousseau and Kant.

In this review I cannot, of course, hope to cover very thoroughly matters that will no doubt be debated in the literature for many years to come. I shall, however, take up a number of points that seem to me to merit special attention. The section immediately following considers Rawls' view of the relation between justice and other social values and so attempts to outline the framework within which his theory operates. In subsequent portions of the review I shall try to make clear the nature of his contractarian theory of justification and to show how Rawls argues for his description of what he calls the "initial position" of his rational contractors. I then discuss two of the basic principles in his conception of justice—the "equal liberty prin-

* I am indebted to Michael Bratman, David Malament, and Adina Schwartz for stimulating discussions of Rawls' book. Each has contributed to some of the arguments in this review, though each would disagree with various other ones.

† Professor of Philosophy, Rockefeller University. B.S., 1949; M.S., 1951; Ph.D., 1957, University of Michigan.

1. J. RAWLS, *A THEORY OF JUSTICE* 172 (1971) [hereinafter cited to page number only].

2. P. 121.

principle" and the "difference principle." In the final section of the review I shall consider Rawls' account of how the parties in the initial position use their basic principles of justice to design various political, social and economic institutions.

I

I find some confusion at the start in understanding precisely what question Rawls' contractarian theory is intended to answer. In particular, I am not sure whether he is proposing an ultimate standard for determining when actions and social arrangements are *right*, all things considered, or whether he is proposing a criterion for answering what I take to be the somewhat narrower question of when acts and arrangements are *just*. If there are considerations other than, or in addition to, justice that can enter into a decision about what is right, a demonstration that a policy or practice is just does not necessarily settle the question of "on balance rightness" with finality. Conducibility to liberty, general health and happiness, perfection of character, social progress, and economic growth, I should think, are also considerations that have weight, even if we admit that considerations of justice are those that should have the most weight. It follows that a practice can be right even though to some extent unjust, and that we can sometimes be *justified*, all things considered, in treating some persons to some extent unjustly. Because of these unhappy truths, some philosophical writers³ have resorted to an invented technical term "justicize" meaning "to show something to be just," in contrast with the word "justify" which means "to show something to be right, all things considered." Rawls, however, apparently rejects the line of thought reflected in this usage. It seems to be his view that if one set of social arrangements is more just than another, it is decisively better than the other, and it would be wrong not to choose it if we can. The other considerations that have a bearing on rightness are relevant only when we must choose between arrangements that are equally just.

Rawls' leading rivals, the utilitarians, have also tended over the centuries to obscure the distinction (if there is one) between justice and on-balance rightness. In their writings about justice it is not always clear whether they are proposing a utilitarian analysis of the concept of justice itself, or applying a utilitarian standard of right conduct to specific social problems often at the expense of the dictates

3. Frankena, *The Concept of Social Justice*, in *SOCIAL JUSTICE* 5 (R. Brandt ed. 1962).

of justice. The following are the main possible views about the relations between justice and utility:

1. Individual justice and social utility, properly conceived, can *never* conflict. That is because "justice is a name for certain moral requirements, which, regarded collectively, stand higher in the scale of social utility, and are therefore of more paramount obligation, than any others . . ."⁴
2. Individual justice and social utility do sometimes conflict, and when they do, so much the worse for individual justice.
3. Individual justice and social utility do sometimes conflict, and when they do, so much the worse for social utility.
4. Individual justice and social utility do sometimes conflict, but it is impossible in advance to say that one must always have a stronger claim than the other. These opposing irreconcilable claims can only be "balanced" against each other in the concrete circumstances of their conflict.

Views (1) and (2) are both utilitarian, but they are answers to different questions. The first *analyzes* justice in terms of utility and yields the arbitrary and implausible result that the two can never conflict. The second view, characteristic of Stalinists and other wicked Machiavellians, is more realistic, but "realistic" in the pejorative sense of that term. The third position, which accords with the ancient and honorable tradition of "*fiat iustitia, et pereat mundus*" is probably that of Rawls. What I am uncertain about is whether Rawls' view is to be understood as a direct rival to utilitarian view (1) only or to utilitarian view (2) as well. Most of Rawls' arguments seem to be directed against (1). Many of these show that justice as analyzed by utilitarians endorses arrangements as "just" which are plainly contrary to our natural shared convictions. I find Rawls entirely convincing in these arguments. Utilitarian view (2), however, is wholly consistent with a Rawlsian *analysis* of justice, and argues with some plausibility that utility, while different from justice, is nevertheless superior to it. I do not find the arguments in Rawls' book against this second kind of utilitarianism quite as convincing as those by which Rawls destroys the first kind. I am not myself tempted to the second utilitarian view, however, because it seems to me an extreme exaggeration, and therefore a caricature, of the insight of its proponents. Surely, we would not judge it to be right on balance to choose a

4. J.S. MILL, *Utilitarianism*, in UTILITARIANISM, LIBERTY AND REPRESENTATIVE GOVERNMENT 59 (Everyman's ed. 1910).

slightly more “useful” social practice in preference to a vastly more just one. Individual justice must provide *some* limit to the claims of the general welfare, even though the latter may be considerable.

I am inclined, therefore, to reject utilitarianism as either a misconstruction of the relevance of social utility or an exaggeration of its importance, and to consider the fourth view distinguished above to be the principal rival to Rawls’ theory. Rawls calls theories of this kind “intuitionistic,” because they deny that a plurality of moral principles can be reduced to one ultimate and superior one, and they also deny that “priority rules” can be formulated which assign weights to conflicting considerations so that we can determine in particular cases which has priority. Hence, irreducibly distinct principles must be “weighed” or “balanced” in particular cases of conflict by “intuition.”⁵

Rawls’ own view (*if* we interpret him to be talking about on-balance rightness, or justification, and not only “justicization”) is that there are principles other than justice that have some force, but in respect to them, justice has absolute weight, always deserving priority in cases of conflict. Rawls argues against his less absolutistic opponent in a characteristically undogmatic way. He expresses his respect for intuitionism as a plausible rival, and he concedes that he cannot prove it false. His own theory, however, is urged to be superior to it in at least one way: it has no priority problem. Rawls not only has a rigid priority rule giving justice absolute weight over other moral considerations; he also proposes a rigid ordering scheme among the various parts of his principles of justice themselves. If Rawls’ priority rules are defensible, he claims, then his theory is clearly superior to intuitionism:

The assignment of weights is an essential and not a minor part of a conception of justice. If we cannot explain how these weights are to be determined by reasonable ethical criteria, the means of rational discussion have come to an end. An intuitionist conception of justice [rightness?] is, one might say, but half a conception.⁶

This is fair enough. If the sort of “explanation” Rawls seeks is in principle achievable, then the theory that supplies it carries the day. But there is also the possibility that rigid priority rules are, in the very nature of the case, impossible to formulate. Other things being equal, simplicity is preferable to complexity, but a distorting simplicity

5. See Section 7, pp. 34-41.

6. P. 41.

is worse than none. Existentialist writers have spoken of a "tragic surdity" among equally stringent ethical principles. If they and others are right, the conflict between justice and other social principles may pointedly illustrate Aristotle's remark that we should "look for precision in each class of things just so far as the nature of the subject admits"⁷

Even if precise priority rules between equally valid moral principles are impossible, the results may not be as unacceptable as Rawls suggests, nor indeed contrary to the expectations of sensitive persons experienced in the dilemmas of practical judgment and decision. Rational discussion may indeed come to an end before final agreement is possible, but it might still carry us a good part of the way toward that elusive goal in discussions of social questions as in all other practical contexts. A movie critic follows reason in judging a film by its dramatic tension and flow, its theme or moral, its acting, photography, and music. These components are relevant to its final evaluation in a way in which its length, or the previous experience of its director, is not, and reason can demonstrate such judgments of relevance to the satisfaction of any reasonable person. A skilled critic, moreover, can adduce nearly conclusive evidence for the positive and negative judgments he makes about the relevant elements of the film. Still, in the end, he may have to tote up the film's overall score and *balance* it against incommensurable excellences and defects of other films, if he is to make a judgment of comparative worth. There is no cut and dried formula for assigning numbers, or for deciding whether a well acted poor story rates higher or lower than a poorly acted good story. At best, "priority rules" in such contexts are useful rules of thumb, telling us that in general one kind of element should be weighted more than another kind, but not *how much* more, nor even that it *invariably* weighs more. All of this is well enough understood in cinema criticism. A plurality of standards puts a limit on "the means of rational discussion" short of necessary and universal agreement, but the limit is not all that binding, and it gives no support to the view that film criticism is "all a matter of taste." The same I think is true of moral criticism and judgment. We may not be able to say whether the duty to keep promises or the duty to tell the truth must *always* take precedence in the cases where they cannot both be honored. Nevertheless we can have rational grounds for honoring both duties in the multitudinous cases where they don't conflict, and for arrang-

7. ARISTOTLE, NICOMACHEAN ETHICS 1094b, 25, Bk. I: Ch. 3.

ing our lives in such a way that the occasions for conflict among acknowledged duties is kept to a minimum. Even in actual cases of conflict, the concrete context may make it perfectly obvious which conflicting duty takes precedence in the circumstances. The burden is on Rawls, then, to establish his proposed rigid ordering of justice and social utility in the teeth of skeptical intuitionists (like this writer) and we need not be surprised, much less distressed, if he has failed.

Part of the test of a proposed priority rule in moral philosophy is whether it conforms with our considered confident judgments in specific cases, actual and hypothetical. The intuitive force of Rawls' position is well captured in an example from William James:

If the hypothesis were offered us of a world in which Messrs. Fourier's and Bellamy's and Morris's utopias should be all out-done, and millions kept permanently happy on the one simple condition that a certain lost soul on the far-off edge of things should lead a life of lonely torture, what except a specific and independent sort of emotion can it be which would make us immediately feel, even though an impulse arose within us to clutch at the happiness so offered, how hideous a thing would be its enjoyment when deliberately accepted as the fruit of such a bargain?⁸

The "emotion" to which James refers is the very sense of justice whose deliverances Rawls venerates. Even those of us who are so hardened against its voice that we would make the deal with James' devil would be troubled ever after by the thought that the great net gain in human happiness so achieved was made *unfairly* at the expense of the innocent victim.

Most of us, I feel confident, would join James and Rawls in declining the devil's offer. No conceivable gain in social well being could possibly balance the cosmic injustice done the lost soul. Even utilitarians might be forced by this example to interpret "social utility" in a broad enough way to escape its thrust. If every human soul, as such, has infinite value, they might argue, then moral mathematics will not dictate the sacrifice of one for the sake of the others. The laws of ordinary arithmetic do not apply to infinite magnitudes. The value of one infinite soul would be precisely equal to that of two billion other infinite souls taken together, and therefore the suffering of the sacrificed one would have no less negative value than the per-

8. W. JAMES, *The Moral Philosopher and the Moral Life*, in *ESSAYS IN PRAGMATISM* 68 (Hafner, 1948) [hereinafter cited as JAMES].

fect happiness of the benefited souls had positive value. But this way of interpreting the metaphor of "infinite value" will not completely save the utilitarian from embarrassment, if he is forced to admit that on his principles the decision whether to accept the offer of James' devil is a matter of moral indifference, a "six of one and half dozen of the other."

When we consider examples of a different kind, however, Rawls' absolutism loses plausibility. Suppose that our option is not whether to promote social *gains* unfairly at the expense of sacrificed individuals, but rather whether to prevent social *harms*—widespread suffering, misery, poverty, and ignorance—at the expense of sacrificed individual or class interests. Such options have in fact been faced over and over again by political leaders since the time of the first industrial revolution, and the case for ruthless means to the end of industrial growth, power, and prosperity has often been made with disconcerting persuasiveness. Even our own unrivaled affluence has been said to be an historical vindication of the atrocious conditions imposed on the working masses in the earlier periods of industrial capitalism. But the case was put most forcefully by the defenders of Stalin's ruthlessness in transforming a backward agricultural nation into the world's second industrial power. In Koestler's novel *Darkness at Noon*, the commissar Ivanov "justifies" the murders of millions of kulaks to the prisoner Rubashov:

"For a man with your past," Ivanov went on, "this sudden revulsion against experimenting is rather naive. Every year several million people are killed quite pointlessly by epidemics and other natural catastrophes. And we should shrink from sacrificing a few hundred thousand for the most promising experiment in history? Not to mention the legions of those who die of undernourishment and tuberculosis in coal and quicksilver mines, rice-fields, and cotton plantations. No one takes any notice of them; nobody asks why or what for; but if here we shoot a few hundred thousand objectively harmful people, the humanitarians all over the world foam at the mouth. Yes, we liquidated the parasitic part of the peasantry and let it die of starvation. It was a surgical operation which had to be done once and for all; but in the good old days before the Revolution just as many died in any dry year—only senselessly and pointlessly. The victims of the Yellow River floods in China amount sometimes to hundreds of thousands. Nature is generous in her senseless experiments on mankind. Why should mankind not have the rights to experiment on itself?"

He paused; Rubashov did not answer. He went on: "Have you ever read brochures of an anti-vivisectionist society? They are

shattering and heartbreaking; when one reads how some poor cur which has had its liver cut out, whines and licks his tormentor's hands, one is just as nauseated as you were tonight. But if these people had their say, we would have no serums against cholera, typhoid, or diphtheria . . ."⁹

A sophisticated utilitarian philosopher, of course, need not agree with the commissar Ivanov. He could (and probably would) argue that Stalin's "utilitarianism" was very naive indeed, and that the cruel murders done in its name were done in deplorable hastiness and ignorance of actual causal connections, and through egregious miscalculations of probability that made them self-defeating and unnecessary. Valuing the general welfare does not commit one to ineffective and uneconomical means to achieve it, means which spew misery and distrust as side effects. In short, there is no utilitarian justification of idle or uncertain *experiments* with high initial costs in suffering. We do not sacrifice human beings, even volunteers, on the *mere chance* that we may find the cure for cancer, or the like.

Wholly utilitarian policies, however, are not necessarily naive and uninformed, and in any case as philosophers we need only to resort to plausible hypothetical examples of a sophisticated kind. Suppose then that political leaders *know* what means are necessary to avert or eliminate a horrible widespread evil, and that these means require that the interests of innocent persons be sacrificed for the sake of the remainder, as in the many examples in criminal law books of "necessity" or "forced choice of the lesser evil." Two variants of the problem can be distinguished. In the first, those who are sacrificed would be lost in any case, so that the choice is between saving some or saving none at all.¹⁰ In that kind of situation, there is no clear conflict between justice and the greater good. In the other variant, however, the interests of some persons are sacrificed for the sake of the interests of others, leading to a vast net improvement in social well-being, in circumstances that do permit an alternative in which the minority interests would not be sacrificed and social misery not remedied. A case in point would be the option of a dictator in a backward country to sacrifice the liberties of many of his subjects in order to accelerate the economic development held necessary for the elimination of wide-

9. A. KOESTLER, DARKNESS AT NOON 161-62 (1941).

10. See Stephen's example in his discussion of Regina v. Dudley and Stephens, 14 Q.B.D. 273: "Several men are roped together on the Alps. They slip and the weight of the whole party is thrown on one, who cuts the rope in order to save himself. Here the question is not whether some shall die but whether one shall live." J. STEPHEN, DIGEST OF THE CRIMINAL LAW n.4 at 24 (4th ed. 1887).

spread misery in the present and future generations. It is by no means evident to me that the morally right policy in such circumstances is *always* to "Let justice be done though the heavens fall and the masses perish." Nor is it self-evident that the correct policy must always be to sacrifice some for the sake of the rest. An intuitionist is a philosopher who cannot commit himself in advance to a single moral principle that can be a trump card in all cases of this kind. There are, after all, different degrees of injustice, different amounts and types of evils, and different probabilities in our projections of consequences. Many of these factors are mutually incommensurable. To be sure, there are easy cases too, and even well supported general maxims, such as that directing us to take individual justice more seriously, by and large, than social utility, especially given our propensity to miscalculate and overestimate the latter. Still, there are close cases too, and few sensitive persons will be satisfied with a theory that would represent even the difficult problems as simple by declaring in advance that one type of conflicting consideration must always triumph over the other.

II

So much for Rawls' view of the relation between justice and other social values. His main concern, however, is not with that question at all but with the analysis of the concept of justice and the derivation of its basic normative principles. In the performance of that task, Rawls' treatise, while not without its difficulties, is unexcelled in the philosophical literature. His main rival in this enterprise is utilitarian theory (1) above, and against it Rawls wins a decisive victory.

Lengthy and thorough as this book is, it still is not a complete theory of justice. Rawls explicitly restricts its scope in two ways. First, he is almost exclusively concerned with the justice of basic political and economic *institutions*, as opposed to the justice of individual actions, persons, or policies. "The primary subject of justice is the basic structure of society."¹¹ Second, he is concerned to describe an ideally just basic structure as it would operate if everyone always acted justly and discharged the duties imposed by the basic structure and its derivative rules. Rawls' book then is essentially a treatise in "strict compliance theory" as opposed to "partial compliance theory."¹² Rawls rightly holds that the description of an ideally just basic structure under conditions of strict compliance is the first and most fun-

11. P. 7.

12. P. 8.

damental task of a theory of justice,¹³ but his relative neglect of specific problems of decision and judgment under conditions of only partial compliance will disappoint readers with jurisprudential interests. The occasion for a lawyer's concern with justice is typically when something has gone wrong, when duties are flouted, and rights violated. There is very little discussion in Rawls' book, for example, of the conundrums of criminal justice, or the role of fault in the law of torts. The book's only detailed discussion of a problem in partial compliance theory is its treatment of civil disobedience and conscientious refusal.

There are many elements in Rawls' complete theory, but the most basic ones are those indicated in the following outline. The theory consists of:

- (1) *A contractarian theory of justification.* This is the theory that the only way to justify substantive principles of justice is to demonstrate that they are uniquely "rational," that is, that they are precisely the principles that would be selected by any and all rational persons in an "initial position" in which the basic structure of society is to be designed from scratch.
- (2) *A "philosophically adequate" theory of the conditions that characterize the initial position.*
- (3) *A deduction of the correct basic principles of justice from (1) and (2).* These principles are (in order of priority) —
 - (A) *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.
 - (B) *Second Principle.* Social and economic inequalities are to be arranged so that:
 - (i) they are attached to offices and positions open to all under conditions of fair equality of opportunity, and
 - (ii) they are to the greatest benefit of the least advantaged ("the difference principle").
- (4) *A design of a "basic structure" of institutions and practices that accords with (3).*

My interpretative and critical comments will follow this outline and fill in some of its gaps.

First of all, what precisely is a "contractarian theory of justification"? Many lawyers will be relieved to learn that Rawls does not present a new or detailed theory of the nature of contractual obligation, nor does he base political obligations on some actual, tacit, or fictitious contract agreed to by ourselves or our ancestors. Rather, in

13. P. 9.

this part of his theory, Rawls describes what he takes to be a *test for the truth* of proposed principles of social justice, a test that is necessarily formulated in the subjunctive mood. A proposed principle is correct providing it *would be chosen* over any alternative principle that could be proposed to a group of normally self-interested rational persons, each concerned to advance and protect his own interests and those of his household. All of these persons can be imagined to have gathered together voluntarily in a condition of equal power with the ultimate purpose of designing the institutions that will regulate their future lives. We are also to suppose that each wears a kind of "veil of ignorance" which prevents him from knowing facts about his future condition that could tempt him ("rationally" self-interested being that he is) to base his choice on the desire to promote his own interest to the disadvantage of others. Thus—

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.¹⁴

Nor do these hypothetical contractors ("choosers" would be a less misleading term) even know to which generation in the total history of their society they will belong. We can ascribe to the choosers almost unlimited "general knowledge," including knowledge of the laws and theories of physical and social science. The point of the veil of ignorance barring knowledge of various kinds of particular facts is to "nullify the effects of specific contingencies which put men at odds and tempt them to exploit social and natural circumstances to their own advantage."¹⁵ If the hypothetical choosers are neither rich nor poor, white nor black, male nor female, old nor young—so far as they know—then they *can not* be lobbyists for any particular class interests, and must choose their principles from a more disinterested vantage point.

In what sense are the hypothetical choosers "rational"? Rawls must

14. P. 137.

15. P. 136.

take great pains to keep various kinds of moral biases, tastes, and interests out of his conception of rationality. The more spare his conception of rationality at this stage, the more impressive will be his claim that substantive principles of justice can be *deduced* in a non-circular, non-question-begging way from pure disinterested reason itself. Whether we think Rawls successful in this enterprise or not, we must admit, I think, that his arguments make the clearest and most impressive case thus far for the ancient doctrine that justice is “derived from reason itself,” and that injustice—in the words of Hobbes —“is a kind of absurdity.” The rationality that Rawls attributes to his hypothetical bargainers is much like the spare concept of rationality used in economic theory. A rational person avoids logical errors in his inferences and can choose means appropriate to his ends, but his rationality, as such, commits him to no particular set of ends. His rationality implies only that he has

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.¹⁶

The only particular wants that can be derived from this formal conception of rationality are wants for “primary goods.” These are goods that are useful to any person whatever the particular nature of his goals—rights and liberties, powers and opportunities, income and wealth. It is of the essence of rationality to protect one’s liberties, widen one’s opportunities, and enlarge one’s means for promoting one’s aims “whatever these are.”¹⁷ The final “assumption” Rawls makes about rationality is that it is irrational to cut off one’s nose to spite one’s face, to be so moved by envy of another as to prefer a net loss for oneself to a distribution in which the other has a larger share.¹⁸

The contractarian method is to attempt to demonstrate that *any* rational person so conceived, whatever his actual class, sex, race, age, generation, or ultimate goals, if he is normally self-interested, and wearing the appropriate veil of ignorance, would choose one definite set of principles as the basic ground rules of the society he is to join. A collection of such choosers then would choose those ground rules *unanimously*. According to contractarianism, then, to show that a

16. P. 143.

17. *Id.*

18. See § 80, “The Problem of Envy,” at 530-33.

proposed ground rule is indeed *the* correct or true principle of justice is to *deduce* it from these detailed assumptions about the beliefs and interests of rational men in a definite set of circumstances with a precise list of options open to them. The task of a complete contractarian theory is actually to deduce those ground rules as theorems in a kind of moral geometry.

The advantages of a contractarian theory of justification in the theory of justice (as opposed to the theory of right conduct generally) are impressive. The real point of contractarianism is to capture and conserve the *objectivity* peculiar to judgments of justice and injustice as they are made in ordinary life, public arenas, and courts of law. Judgments of justice are *par excellence* the moral judgments made in a spirit of neutrality, impartiality, impersonality, and disinterestedness. When we make such judgments we are understood to be claiming to speak from an impartial observer's vantage point, not expressing merely personal or sectarian attitudes, nor advancing merely selfish or sectional interests. The language of justice, of course, can be a smokescreen concealing the real motives of the speaker or judge, but it cannot even serve to conceal in this way if it is conventionally understood to be a vehicle for subjective interest or bias. If I say that "*x* is just," I cannot *mean* that *x* serves my interests or the interests of those I favor. Such an interpretation is absolutely ruled out, whatever my real intentions. Just ground rules, whether for a newly designed game,¹⁹ or a newly designed society, are those that give no unfair advantage to anyone, or those that would be designed by an altogether impartial legislator. The contractarian, impressed by this point, finds a more fruitful but equivalent way of making it. Just ground rules, he says, are those that would be chosen in advance by the players themselves by some *fair procedure*. (Hence, the name Rawls applies to his basic conception: "justice as fairness.") The veil of ignorance and the circumstances of the "initial situation" are important parts of the conception of a fair, unbiased, and impartial procedure for agreeing to basic rules.

What are the alternatives to a contractarian theory of justification? Rawls himself cites utilitarianism as the chief rival to his contractarianism,²⁰ but once again the reference to utilitarianism is confusing. The word "utilitarianism" is the name of so many different philosophical

19. The imperfect analogy of a game, however, while useful for some purposes, is quite misleading for others, as Rawls is quick to acknowledge: "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." P. 145.

20. P. vii.

doctrines, answering so many different philosophical questions, and often standing or falling in logical independence of one another, that I would advocate its banishment once and for all from the philosophical vocabulary. I have already distinguished utilitarianism as a standard of right conduct from utilitarianism as a standard of just arrangements. Utilitarianism as an ultimate test of truth for proposed principles of justice would be still less plausible. Nor would it be very useful for the utilitarian philosopher who wishes to establish the truth of a substantive utilitarian principle of justice on objective and independent grounds. If, for example, he attempted to defend as a principle of just economic distribution the doctrine that goods should be distributed so as to maximize the net total (or alternatively, average) wealth even at the cost of impoverishing the least advantaged class to the point of starvation, he could only derive it from a utilitarian justificatory criterion from which it follows tautologically. There would be no great gain in conviction or persuasiveness in that circular test. It would be much like a judicial system in which a trial court and an appellate court were one and the same, so that a single judge would control the course of trials at one level and then uphold his own judicial pronouncements, speaking from a higher level. Such a system of appeals would hardly strengthen the appearance of impartial objectivity in the system as a whole. Most proponents of utilitarian principles of justice would presumably prefer to appeal to a highest tribunal of reason that is not itself utilitarian, so that their favorite principles could receive a stamp of approval other than their own. In the context of Rawls' book, therefore, "utilitarianism" should be understood as one of the numerous proposed ground rules considered for adoption by the parties in the initial situation.²¹ Rawls' contractarianism, rather than being a direct rival to utilitarianism, is a kind of ultimate tribunal to which the parties appeal in deciding among alternative principles. It is for Rawls a kind of test which his own principles of justice pass, and utilitarianism, in a properly narrow sense, fails.

It should also be emphasized that there are thousands of *contractarian* alternatives to Rawls' own theory. Other contractarian theories can be generated by altering any of the components of Rawls' special conception of rationality, or by restricting the choosing parties to individuals (excluding families), or by permitting pre-existent corporate associations to be choosers, or by introducing special assumptions about

21. As Rawls himself does throughout most of his book. See especially pp. 122-26.

scarcity or abundance of available economic resources (Rawls, following Hume, assumes "moderate scarcity"), or by raising or lowering to some degree the veil of ignorance, by increasing or decreasing the altruism or selfishness of the bargainers, and in various other ways. Rawls' defenses of each element in his description of the initial position are subtle and thorough. They must be so, since they constitute a crucially important section of his total theory. But insofar as we reject any part of Rawls' account of the parties and their circumstances in the initial position, we revise the premises from which the correct principles of social justice are to follow as moral theorems. Thus, an alternative theory of the initial position will yield different ground rules for the just society.

III

How does Rawls argue for his description of the initial position? He makes his argument by appealing to considerations of simplicity and by selecting only the weakest assumptions so as to enhance the aura of disinterested objectivity so essential to those principles if they are to be called principles of *justice*. But Rawls also appeals at numerous places to a class of considerations which are likely to disconcert the reader who does not fully understand his methodology; for these considerations, at first sight, appear to indict the basic argument of his book, taken as a whole, for *circularity*. Just as his description of the initial position serves as a premise in his deduction of the principles of justice, so his antecedent and unargued conception of those principles serves as a test for the various parts of his description of the initial position. In other words, Rawls begins his description of the rational parties and their circumstances in the initial position already possessed of an intuitive idea of the principles of justice he wishes to justify, and then considers what those initial conditions must be like in order to generate the principles he already knows to be correct. But there is nothing sneaky or sleight of hand about this procedure. Rawls confesses it openly in numerous places and is unembarrassed by it. Thus, to cite just one example, he writes that "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."²² Rawls' procedure *is* circular, but not, I believe, in any vicious or invalidating sense. To understand why this is so, we must turn to the topic of methodology in moral philosophy.

22. P. 141.

Rawls shares a conception of the modes of argument available to the moral philosopher with a growing number of other writers.²³ Perhaps the least misleading label for the conception would be "the coherence theory." The writers to whom I refer have been disillusioned with various traditional modes of argument in moral philosophy. They do not believe it possible to base an ethical system on self-evident moral first principles, or on direct intuitive insight into, or rational apprehension of, a uniquely moral realm of truth. Nor do they think it possible to deduce moral first principles from statements of fact, making no challengeable moral assumptions along the way. On the other hand, these writers are not willing to deprive general principles of their usual role in arguments for relatively specific maxims and judgments. General principles and factual premises do entail specific moral judgments, they admit, but the most suitable general principles, they insist, are those that summarize and are supported by the specific moral judgments in which we have the most confidence. We justify specific moral judgments, on their view, by deriving them from general principles, and the latter are supported in turn by a demonstration that the right moral judgments (*other* moral judgments) follow from them. This may be circular, but it is unavoidably and non-viciously so.

Most coherence theories of ethical justification contain both a "logic of discovery" and a "logic of justification." The former governs the reasonings of a single person, trying by himself to reach a judgment about what is right, or just, or good. The latter provides basic rules for the rational resolution of disagreements between individuals. Both procedures involve a large element of trial and error and muddling through, and neither starts altogether from scratch in a kind of moral vacuum. Coherence theorists insist that there must already be some moral conviction, however weak and tentative, before we can arrive at or "prove" other judgments. The individual who is trying to make up his own mind begins with the set of moral beliefs and sentiments he already has, for better or worse. If he is reasonably open-minded he will regard all of these as merely tentative and subject

23. Among the earliest antecedents of this view is that of ARISTOTLE, *Nicomachean Ethics* 1094b, 12-1096a, 10, Bk. I: ch. 3-4. Recent statements of it include A. EWING, *Ethics* 1-15 (1962); JAMES, *What Pragmatism Means*, in JAMES, *supra* note 8, at 141; M. WHITE, *Toward Reunion in Philosophy* 251-63 (1956); Feinberg, *Comments*, in *Society: Revolution and Reform* 19-32 (R. Grimm and A. MacKay eds. 1971). For models from the philosophy of science, see N. GOODMAN, *Fact, Fiction and Forecast* 65-66 (1955), and W.V. QUINE, *Two Dogmas of Empiricism*, in *From a Logical Point of View* 20 (2d ed. rev. 1964). For applications to the philosophy of law, see Sartorius, *The Justification of the Judicial Decision*, 78 *Ethics* 171 (1968), and Scheffler, *On Justification and Commitment*, 51 *J. of Phil.* 180 (1954).

to future correction, but some will surely seem more tentative than others. He places his beliefs then in an order of conviction. Some of them can easily be revised or dispensed with, but at the other extreme, there are some that he would be loath ever to give up. He then reaches the level of general principle by extracting from his most confident convictions their apparent rationale. Then he applies his newly discovered principle to the more difficult cases and decides them as the general principle directs. All goes well until his general principle directs him to a judgment in a particular case that runs counter to one of his most confident convictions. Then there is a crisis and something will have to give. The problem for practical reason in such a crisis is to reformulate the general principle (often by appending an exceptive clause) in such a way that it still faithfully summarizes the whole body of one's singular convictions and no longer yields an unacceptable result. In other cases, where a deeply entrenched and often tested general principle conflicts with a specific moral belief of relatively minor conviction, it is the latter which must be changed in the interest of internal harmony. Moral experience constantly throws up new and unanticipated problems so that the task of "making everything come out even" is never finished.

The coherence logic of justification muddles through in similar ways. We take as our simplest model here two individuals who are in disagreement over a moral question. (A variant is the case where one is trying to justify one of his beliefs to a skeptic who may have no belief of his own on the matter, one way or the other.) If there is no common ground of moral conviction whatever between the two individuals, either at the level of general principle or the level of singular judgment, then the game is over before it begins. Logic cannot build castles on air. It is always a reasonable assumption, however, that two individuals in ethical disagreement over one question can find other matters on which they are in solid agreement. In order for one individual to "prove" that the other is wrong in a given belief he must *use* those beliefs and principles of the other party which he himself shares. He must show the other party that the belief he is currently defending is inconsistent with one of his own principles, or if the subject of dispute is one of the other party's principles, that it would commit him to a specific judgment which he would be embarrassed to make. The way to convince another party of the truth of a general moral principle is to show him that it summarizes and renders coherent *his own actual convictions* better than any alternative principle.

The coherence theory is hardly new as it appears in Rawls' book,

but Rawls gives it an unusually clear and elegant (and to me convincing) formulation. It is plain, I think, that the final court of appeal, or source of evidence, for his whole normative theory (the tenets numbered (2), (3), and (4) in the outline above) is the set of confident convictions about particular questions of justice and injustice that the author presumably shares with his readers, for example "that religious intolerance and racial discrimination are unjust."²⁴ There are other questions of justice which are much more difficult or to which we have given less thought. Even these, however, are part of the test for proposed general principles, for those principles must not only fit our relatively fixed convictions, they must also help us extend those convictions in "an acceptable way."²⁵

In describing the conditions of the initial situation, then, Rawls is admittedly tampering with the premises from which his proposed principles of justice are to follow with an eye to secure as good a fit between those principles and our antecedent considered judgments as possible. And so, at this stage of the theory, "we work from both ends,"²⁶ at one end describing the conditions of our hypothetical initial position, and at the other, shaping those general principles (that we have extracted from our antecedent judgments) which will then be deduced from those conditions. There are bound to be discrepancies, at first, between the conditions (premises), the principles (to be deduced), and the antecedent judgments (to be matched). In that case, we have the option between making revisions at one end or the other:

By going back and forth, sometimes altering the conditions of the contractual circumstances, at others withdrawing our judgments and conforming them to principle, I assume that eventually we shall find a description of the initial situation that both expresses reasonable conditions and yields principles which match our considered judgments duly pruned and adjusted. This state of affairs I refer to as reflective equilibrium. It is an equilibrium because at last our principles and judgments coincide; and it is reflective since we know to what principles our judgments conform and the premises of their derivation.²⁷

IV

I turn now to Rawls' basic principles of justice themselves (number (3) in the outline above) with a greater interest, for the moment, in

24. P. 19.

25. P. 391.

26. P. 20.

27. *Id.*

their match-ups with considered judgments than with their artificial derivation from rational choices in an initial position. The basic principles further articulate the traditional doctrine that equals are to be treated equally unless there is sufficient reason to treat them differently. Rawls' fundamental thesis is that the only sufficient reason for departing from equality in the distribution of primary social goods—the only reason acceptable to rational bargainers wearing veils of ignorance in the initial position—is that an unequal distribution would be to *everybody's advantage*. This basic idea when it applies to the distribution of political rights and duties is expressed in the First Principle ((3) (A) in the outline) which we can call the "equal liberty principle." Insofar as the basic idea is used to govern the distribution of social and economic advantages it is expressed in the Second Principle ((3) (B) in the outline). The two parts of the Second Principle can be called, respectively, the "open offices principle" and the "difference principle."

I shall consider the equal liberty principle first. The primary goods whose distribution is to be regulated by this principle are the basic liberties of citizenship of the sort conferred by the American Bill of Rights. These include the right to vote and run for public office, the right to free speech and assembly, the right to hold personal property, the right to freedom from arbitrary arrest, to habeas corpus, and other due process protections. The equal liberty principle decrees that every citizen, without exception, is to have all of these rights and to have them equally. That is to say, in respect to citizenship there is to be only one legal status—full membership in the state. For a party in the initial position to agree to inequality in the distribution of these basic liberties (that is, to a multi-status society) he would have to be convinced that the inequality would be in everyone's interests, even the interests of those deprived of full membership. (For all the party in the initial condition knows, he himself could end up in that deprived class.) Rawls is very firm in insisting that inequality of basic rights *cannot* be in the interests of *everyone*.

Rawls' discussion of liberty is thorough and sophisticated, but it seems to this reviewer to suffer from a somewhat vacuous formalism. First of all, the formulation of the principle itself suggests that a large number of different systems of rules, including some that jar with our "considered beliefs in reflective equilibrium" could in principle satisfy it. "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."²⁸ A system in which everyone had the absolute right to move his hands freely through space as he wishes would satisfy that formula every bit as well as a system in which everyone had the absolute right not to be punched in the nose. In the latter system the "inviolable nose interest" would prevail over the "free hand motion interest"; in the former, the priorities would be reversed. Yet both rules could satisfy Rawls' principle. In the one case, citizens are given as much hand motion liberty as is compatible with a like amount of that liberty for all; in the other system, all citizens are granted as much nose protection as is compatible with a like amount of *that* liberty for all. Perhaps Rawls would wish to say that the one liberty is *more basic* than the other, but in that case we need some criterion of "basicness," and I find none in Rawls' book. The problem is especially urgent (and especially interesting to lawyers) in respect to the line between "the right to know" and (others') rights not to be known about in certain respects, between the right to the free exercise of one's religion and (others') rights to be protected from noise or disease, between liberty of speech and (others') rights not to be abused or defamed.

Over and over again, Rawls insists that "the limitation of liberty is justified only when it is necessary for liberty itself"²⁹ or "to avoid . . . an even greater loss of liberty."³⁰ The primary purpose of these formulas is to express the priority of the First Principle over the Second, that is, to insist that "greater economic and social benefits are not a sufficient reason for accepting less than an equal liberty."³¹ That is a very important and controversial substantive claim, which I shall consider below, but the language to which it gives rise—that liberty can be restricted only for the sake of liberty—can be seriously misleading. It suggests that liberty, or at least "basic liberty," is all of a piece, and it conceals from us the true nature of our dilemma when we must draw lines precisely between one liberty and another. No one has a single undifferentiated interest in liberty as such. Rather we have interests in speaking out, debating, voting, practicing our religions, reading books and newspapers, petitioning for grievances, traveling from one place to another, being alone and unobserved with ourselves or others, and many more. Sometimes one person's interest of one kind must be interfered with for the sake of other persons' interests

28. P. 302.

29. P. 215.

30. P. 214 *passim*.

31. P. 207.

of the same or a different kind. The justification for such interferences is not that they are necessary to prevent even greater depletions in the overall supply of some homogenous thing called "liberty," but rather that some interests are more important, more worth protecting, than others. The importance of a given interest must itself be established by some standard other than mere "conducibility to liberty" if it is to be established at all.

I have one other quarrel with Rawls' excellent discussion of liberty. He wisely follows Isaiah Berlin³² in distinguishing between the liberties conferred and protected by the state and the contribution such liberties make to their possessors' capacities to advance their ends. Rawls formulates the distinction as that between *liberty* and the *worth of liberty*. Clearly, the worth of a system of equal basic liberties to all who possess them will not itself be equal so long as there are inequalities in authority and wealth among citizens. Some social and economic inequalities will be to everyone's advantage, and therefore will be justified by Rawls' Second Principle. It follows therefore that differences in the *worth* of the "equal" liberties will be justified too. Rawls is adamant against proposals to compensate those whose basic liberty is worth less by granting them still greater amounts of basic liberty, even in the ideally just society he sketches, where we need not be concerned with practical difficulties. His reason is that that kind of compensation, which would be in strict violation of the equal liberty principle, is unnecessary. That is because, in a certain sense, the lesser worth of liberty in the ideal state is already compensated for "since the capacity of the less fortunate members of society to achieve their aims would be even less were they not to accept the existing inequalities whenever the difference principle is satisfied."³³ But this misses the point. It is a good reason, perhaps, against offsetting the lesser worth of some basic liberties with economic grants even when that would violate the difference principle, for *ex hypothesi*, such subsidies would leave the beneficiaries of our compensation worse off still. But it is no argument against making the liberties themselves unequal in order to offset inequalities in their "per unit worth." I have in mind such schemes as giving additional votes in elections to socially and economically disadvantaged classes. Practical difficulties aside, I should think that precisely such measures are called for by Rawls' equalitarian basic principle.

I shall pass quickly over the open offices principle and linger only

32. I. BERLIN, *Two Concepts of Liberty*, in *FOUR ESSAYS ON LIBERTY* 118-72 (1969).

33. P. 204.

briefly over the “difference principle.” The latter is the basic determinant of economic justice in Rawls’ system and is certain to receive vast amounts of close attention from critics elsewhere. The difference principle gives expression to a very simple intuition. A very good reason, everyone should admit, for consenting to an unequal distribution of goods is that every party to the distribution, even the party who is to receive the smallest share, will thereby receive a larger share than he would have if the distribution were strictly equal. The rules that determine the way an economic system distributes its output, as both capitalists and socialists agree, can causally influence the size of the output produced. The difference principle derives from that acknowledged fact and is familiar to all of us as at least *one* of the principles to which both liberals and conservatives appeal when they debate the justice of economic inequalities between classes. A standard conservative defense of inequality is that it derives from a system that produces more even for those who get smaller shares. The standard liberal indictment of the inequalities they perceive to be unfair is that they are *unnecessary*, that is, they do not make even the smaller shareholders better off. We can also agree with Rawls (but there is no space here to examine his arguments) that a merely utilitarian defense of inequality, to the effect that it is a necessary means to greater total wealth or greater average wealth, though not necessarily to greater wealth for the least advantaged, would be much less likely to assuage the sense of injustice when it has been inflamed by the awareness of inequalities of wealth and income. The difference principle, then, is *a* powerful reason (as its utilitarian rivals are not) for the justicization of inequality. The much more difficult and controversial question is whether it is the *only* reason that can play that role.

At this point it is important to remind the reader that the basic principles of justice for Rawls are not meant to determine directly the justice of particular actions, policies, or allocations, so much as to govern the design of the basic institutions of society. This is a rule-contractarian theory. “It is a mistake . . . to require that every change, considered as a single transaction in isolation, be in itself just.”³⁴ A given economic transfer, in particular, is sufficiently just, or just in the only intelligible or relevant sense if it follows from the fair operation of complex social machinery which has been designed to accord with the principles of justice that would be chosen in the initial posi-

34. Pp. 87-88.

tion. "A distribution cannot be judged in isolation from the system of which it is the outcome" ³⁵

His emphasis on rules, institutions, and systems leads Rawls to make a somewhat exaggerated claim, at least as measured by its "fit" with our ordinary judgments in reflective equilibrium. The basic structure, he says, and in particular the basic economic system, is a case of "pure procedural justice." Rawls explains what he means by reference to a system of gambling. If we are betting on races, shooting craps, or playing poker, and we play by fair rules honestly administered and scrupulously followed, then whoever wins, the outcome is just. That is because "there is no independent criterion for the right result: instead there is a correct or fair procedure such that the outcome is likewise correct or fair, whatever it is, provided that the procedure has been properly followed."³⁶ In contrast, a criminal trial exemplifies what Rawls calls "imperfect procedural justice."³⁷ There we do have an independent criterion for just outcomes: a verdict is just when the guilty are convicted and the innocent acquitted.³⁸ A system of procedural rules for court trials is fair when it brings about just results in a greater proportion of cases than any alternative, but no system is foolproof.

Rawls is surely right when he points out strong similarities between distributive justice and pure procedural justice, but I should think that economic justice has some elements of imperfect procedural justice too. In some cases, I should think, we can have more confidence that a given allocation is unjust than we have in the fairness of the system that produced it, even though we are confident of the latter too. That implies that we do have some independent notion of a just outcome. In cases of that kind, it may be best not to tamper with the system as a whole, and find means instead to remedy directly the unjust result produced by its operation. So thought Aristotle,³⁹ anyway, and his thought provides the rationale for such corrective institutions as equity courts and ombudsman's offices.

The difference principle surely states a good general reason for departing from strict equality in the distribution of goods, but there

35. P. 88.

36. P. 86.

37. P. 85.

38. This example will do as a rough illustration of imperfect procedural justice, but it oversimplifies the criminal law as it functions in our society. It is sometimes a just outcome when morally culpable defendants are acquitted, as when they can plead the defense of entrapment or when an exclusionary rule prevents the use of incriminating evidence against him.

39. ARISTOTLE, *NICOMACHEAN ETHICS* 1137a31-1138a41, Bk. V: Ch. 10.

is some doubt in my mind whether it always states a sufficient reason, and also some doubt whether it states a necessary condition. If a reasonable person has to choose between two states of affairs, one of which gives him and everyone else in his group (equally) two pieces of pie, the other giving him three pieces and other persons as many as nine or ten, surely he will prefer the latter. If he is a pie lover, he will prefer, *ceteris paribus*, having three pieces to having only two. It would indeed be contrary to reason to prefer a lesser to a greater good for oneself on the sole ground that the greater good would be part of a scheme in which others get greater goods still. (To be so influenced by "envy," Rawls rightly insists, would be irrational.) But if the choice is between two pieces for oneself and all the others, and three for oneself with *ten thousand* for all the others, the differentials are so great that even a generally non-envious reasonable person might be expected to prefer the smaller portion. Several variables seem to be involved: the gap between the smallest and the largest, the gap between the smallest and the second smallest, the gap between the smallest and the average, and so on. My point is that we can fill in the variables in certain extreme ways that provide us with examples (wildly hypothetical ones, to be sure) of distribution patterns that accord with the difference principle, but violate "our considered judgments in reflective equilibrium." The difference principle cites a generally plausible reason for departures from equality that seems to apply to all but extreme cases. The extreme cases show, however, that it is not *always* a *sufficient* reason, and that the narrowing of allocative gaps, as such, has *some* legitimate weight in our ordinary deliberations. Suppose, moreover, that a technological innovation makes possible a substantial gain in wealth or income for those who are already well off, at the cost of only a tiny loss, say a penny a year, for members of the poorer classes. The difference principle would rule out the new allocation as unjust, a judgment which is far from intuitively obvious. That example, hypothetical and unlikely as it is, throws into doubt the claim that the difference principle represents our ordinary judgments in stating a necessary condition for justicized deviations from equality.

Rawls anticipated both of these objections and replied to them as follows:

Part of the answer is that the difference principle is not intended to apply to such abstract possibilities The possibilities which the objection envisages cannot arise in real cases; the feasible set is so restricted that they are excluded. The reason for this is that the

two principles are tied together as one conception of justice which applies to the basic structure of society as a whole. The operation of the principles of equal liberty and open positions prevents these contingencies from occurring While nothing guarantees that inequalities will not be significant, there is a persistent tendency for them to be leveled down by the increasing availability of educated talent and ever widening opportunities.⁴⁰

Perhaps if I understood the logical architecture of Rawls' whole theory a little better, I would know better how to gauge the cogency of this reply. The form of the objection is familiar in moral philosophy. A hypothetical example is presented to discredit a general principle by showing that the principle as applied to it would yield a counter-intuitive result. To the rejoinder that the example is *outré* and unlikely, the counter-rejoinder is usually advanced that if a principle purports to be a criterion for the application of a term to any logically possible situation, it is sufficient for its refutation to provide a logically possible example for which its indicated application would be counter-intuitive. Rawls follows that pattern himself when he is on the offensive against utilitarianism. He argues, for example, that the parties in the initial position would not choose a utilitarian principle for the distribution of their liberties because that principle would authorize the restriction of an individual's liberty just in case this would lead to a greater net balance of utility.⁴¹ To the natural utilitarian rejoinder that given the "general facts of social life . . . the computation of advantages never justifies such limitations," Rawls replies: "But even if the parties were persuaded by this, they might as well guarantee their freedom straightaway by adopting the principle of equal liberty. There is nothing gained by not doing so, and to the extent that the outcome of the actuarial calculation is unclear a great deal may be lost."⁴² The utilitarian, I think, should be subdued by Rawls' cogent reply, but he might well leap to his feet again with a charge of *tu quoque* when Rawls defends the difference principle from "unlikely" or "factually impossible" counterexamples in the way described above.

I come finally to the priority rules which Rawls proposes for the ordering of his two (actually three) principles of justice. The ordering between the principles is to be strictly "serial" or "lexical." He explains that

40. Pp. 157-58.

41. P. 207.

42. *Id.*

[t]his is an order which requires us to satisfy the first principle in the ordering before we can move on to the second, the second before we consider the third, and so on. A principle does not come into play until those previous to it are either fully met or do not apply. A serial ordering avoids, then, having to balance principles at all; those earlier in the ordering have an absolute weight, so to speak, with respect to later ones, and hold without exception.⁴³

Thus, the equal liberty principle cannot be compromised in the slightest for the sake of more equal opportunity to occupy social positions or more equal distribution of economic goods; nor can the open offices principle be departed from at any point for the sake of economic equality. Rawls gives detailed and persuasive arguments for these rankings,⁴⁴ and here and there he gives modest disclaimers cautioning us not to take them too seriously in a non-ideal society. But I do not understand why he need make his ordering for ideal situations so very rigid, or how he can expect us to get along altogether without recourse to “balancing.” (I have already argued that his formulation of the equal liberty principle misleadingly conceals the need to *balance interests* if it is to be applied consistently in the real world.) Again, his theory seems to invite hypothetical counterexamples, some of which, in this case, are not terribly “unlikely.” Is it necessarily unjust, for example, for a government in certain circumstances to concentrate on elimination of extreme economic inequalities between the rich and the starving even at the cost of a modest or temporary suspension of some of the lesser basic liberties? (I have in mind the expropriation of personal property without compensation.) No doubt middle class refugees from the tyrannies of Mao and Castro feel that they have been treated not only harshly but unfairly, and Rawls’ theory gives them solid support. But apologists for Mao and Castro sometimes reply that it would have been unfair (and unfair by the difference principle) to the starving and backward masses for Mao and Castro not to have expropriated the middle class and put it to work. When purely economic disparities between classes are *extreme* they begin to exert a claim on our judgments approaching even that of our basic liberties of thought, privacy, and travel. Is it not possible, in the Chinese and Cuban cases, that justice sits in both pans of the scale, even if its weight is more decisively in one than the other? So at least it seems to the “intuitionist” in the theory of justice—a

43. P. 43.

44. Pp. 243-51, 298-303.

philosopher who resembles Rawls in many important ways but who lacks Rawls' faith in serialized priority rules. Rawls himself says: "At some point the priority of rules for non-ideal cases will fail; and indeed, we may be able to find no satisfactory answer at all. But we must try to postpone the day of reckoning as long as possible, and try to arrange society so that it never comes."⁴⁵ Ah, but that is what the intuitionist says too.

V

The parties in the initial position have agreed on the basic principles of justice for use in the construction of a genuinely cooperative society, but their main work still lies ahead of them. They must now proceed to use their common principles in the design of those political, social, and economic institutions that will most closely accord with these principles. Rawls describes their deliberations step by step in a narrative that is actually dramatic, as he sketches for the reader his own vision of the perfectly just society grounded at every crucial point by a contractarian principle of justification.

The first task of the newly cooperative parties is to design a political constitution. They become delegates to a constitutional convention at which each of them, as a disinterested rational chooser firmly committed to the principles of justice adopted at the previous stage, will agree on a fair system for making, administering, and applying laws, and especially a system for coping with the inevitable political disagreements that will arise when citizens become aware of their actual interests as fully fleshed-out human beings. At this stage the veil of ignorance is partially lifted, and the parties "now know the relevant general facts about their society, that is, its natural circumstances and resources, its level of economic advance and political culture, and so on."⁴⁶ Next comes the stage at which the parties must assess the justice of proposed laws and policies. They are now actual legislators charged with further, more precise, tasks of social construction, within the limits permitted by their newly adopted just constitution. They are still blind to the particulars of their own interests and identities, but by this stage, their impartial disinterest is no guarantee of unanimous agreements. Now they must debate speculative economic and sociological doctrines that bear on legislation, and we can expect reasonable differences of opinion. At the final stage, where rules are applied by administrators and judges and followed by citizens generally, every-

45. P. 303.

46. P. 197.

one's veil of ignorance is completely lifted. Now the hypothetical rational choosers of the initial position have become full blooded citizens, and questions of justice arising among them are to be settled by applying a full system of rules to persons with determinate characteristics in determinate circumstances. Very few questions at this level will require for their resolution more than lower level rules applied to specific facts; that is, justice is very often "purely procedural" and only rarely calls for an application of constitutional rules or the basic principles of justice themselves. Rawls is, for the most part, not an act-contractarian, but a rule-contractarian. (An important exception to this is his subtle treatment of civil disobedience and conscientious refusal, which I cannot discuss here.) "Thus on many questions of social and economic policy we must fall back upon a notion of quasi-pure procedural justice: laws and policies are just provided that they lie within the allowed range, and the legislature, in ways authorized by a just constitution, has in fact enacted them."⁴⁷

I shall not attempt to do justice here to the remarkably detailed sketch of basic institutions which is Rawls' final blueprint for the ideally just society. In certain respects his whole method of rational reconstruction reminds this reviewer of the project of the methodological skeptic, René Descartes, who doubted everything that could be doubted until, freed of dubious presuppositions, he could start from scratch the task of reasoning his way back to solid conclusions about the world. An amazing thing about Descartes' results is that almost everything he doubted in the beginning he "demonstrated" back into existence in the end (but better understood). Similarly, Rawls, who starts with only some abstract rational choosers in a state of nature, plus a handful of assumptions about rationality, concludes by demonstrating the ideal justice of a political and social system very much like our own, one with a written constitution, a bill of absolute rights, a commitment within limits to majority rule, to the rule of law, and to fair equal opportunity (especially educational opportunity), a conservation and savings policy, and a goal of economic equality that includes a test for permissible inequalities acceptable to liberals and conservatives alike. The picture that comes into focus is not that of our blemished actual society, but rather of that society purged of its conspicuous injustices, and true to its noblest ideals. To produce such a picture with such detailed clarity is as practical a service as a philosopher can perform for his fellow citizens, and this citizen, for one, is duly grateful.

47. P. 201.