

RECONCILING LIBERTY AND EQUALITY: JUSTICE AS FAIRNESS

John Rawls's *A Theory of Justice* presents a rich, original, and ambitious theory of justice called "justice as fairness": it is, by general agreement, the most important work about justice in the past century. That of course does not make Rawls's theory right, but his work is of particularly great depth and compelling interest. To be sure, *A Theory of Justice* is also a very demanding work, but it repays all the attention you give it.

Rawls's basic concern might be put this way: What is the most reasonable conception of justice for a society of free and equal persons? What principles should our society meet, if it is to be fair to persons conceived of as free and equal: both conceived in liberty and dedicated to the proposition that all men are created equal? In particular, should it be utilitarian, libertarian, a less liberal egalitarian society, a less egalitarian liberal society? *Theory* offers a three-part answer to this question: first, Rawls presents two principles of justice, and offers an argument for those principles according to which the members of society would choose them in an "original position" behind a "veil of ignorance," which hides all information about their social position and natural endowments. Then second, he presents a sketch of a society that operates on these principles, to show that they are realistic. And he argues, finally, that a just society—just by the lights of justice as fairness—would be stable in part because living in a just society and having a sense of justice guided by his principles is good for those who live in the society.

With time and world enough, we would discuss the entire view, but limits being what they are, we will focus on the first part: the presentation of the principles and the arguments for them, starting with the two principles of justice.

1. Two Principles of Justice

Rawls calls his first principle “the principle of equal basic liberties.” It says that each citizen has an equal right to the most extensive system of equal basic personal and political liberties compatible with a similar system of liberties for others. That may sound like choice-based libertarianism. But where Hayek and Friedman endorse a basic right to liberty as such, Rawls’ principle of liberty requires stringent protections for certain specific liberties—of thought and conscience; political liberties (rights of participation); liberty of association; liberty and integrity of the person; and rights and liberties associated with the rule of law.

The principle of equal basic liberties also includes a strong requirement of political equality. Political liberty is to be assured a *fair value*: chances to hold office and to exercise political influence ought to be independent of socio-economic position, that people who are equally motivated and equally able ought to have equal chance for political influence. This principle has large implications for the way we finance elections: we will return to this issue after we discuss Rawls.

Rawls says that his principles “express an egalitarian conception of justice,”¹ and you can see why when you consider the second principle. Thus his second principle of justice states that socio-economic inequalities must satisfy

two conditions. First, legitimate inequalities must attach to offices and positions open to all under conditions of *fair equality of opportunity*. Here we have the strong requirement of starting-gate equality: in particular, people who are equally talented and motivated must have equal chances to attain desirable positions; a person's fate in life should not depend on the social circumstances of their birth and upbringing.

Second, the difference principle states that socio-economic inequalities, to the extent that there are such, must be to the greatest benefit of the least advantaged members of society. To understand how this works, it is essential to keep in mind that we are not assuming a fixed economic pie: so if inequalities foster economic growth, the results can make everyone better off. Thus consider two cases: compensation and incentives. Someone might legitimately be paid more than someone else because the higher income *compensates* for expensive training and education that enable the person to take on socially desirable tasks; or inequalities might make sense as *incentives* encouraging people to take on tasks they would otherwise be unable or simply unwilling to take on. According to the difference principle such inequalities may be perfectly just, but they must be no greater than necessary to maximally benefit the least advantaged: thus if someone needs an incentive to take on a job that is socially beneficial, the incentive is perfectly just, provided that it is no larger than necessary to get the person to do the job. The ordinary workings of labor markets are not likely to satisfy this principle, because those workings reward people with scarce talents. So tax rates and income transfers must be adjusted with the aim of ensuring that

any increase in tax rates, for example, would worsen the conditions of the least advantaged and any decrease would do so as well.

Justice as fairness, with its two principles of justice, is an alternative to utilitarianism and libertarianism. An alternative to utilitarianism, because the right to equal basic liberties and distributive fairness take priority over maximizing general welfare. Particularly because of this emphasis on the priority of liberty, *justice as fairness* is squarely in the liberal tradition of political thought. But justice as fairness is also an alternative to libertarianism and the classical liberalism of Locke and Smith, Friedman and Hayek, because justice does not require stringent protection for liberty as such, but for basic liberties; and because justice requires—in the second principle—regulation of choice in the name of equality.

2. What, more exactly, do the principles say?

Before turning to Rawls's arguments for the principles, I want to make three further comments on the content of the principles.

First, I said that the protections of the basic liberties are especially stringent: Rawls's equal basic liberty principle has *priority* over the second principle. So Rawls may seem to be endorsing a libertarian view. But his first principle is about specific liberties, not liberty or choices as such. In particular, the market liberties that were the concern of the Lochner Court are not covered by his first principle of justice. So those liberties can be regulated to achieve the aims of the second principle of justice.

So what does the “priority of liberty” come to? It means that justifications for limiting a basic liberty must show how the proposed limit improves the protection of the basic liberties overall. For example, to protect the religious liberty of religious minorities, we might restrict the scope of majority rule by adopting a constitutional right to liberty of conscience that ensures the free exercise of religion: this plausibly counts as a restriction of political liberty, inasmuch as it limits the scope of that liberty. So one basic liberty (political liberty) is restricted to ensure another basic liberty (religious liberty). But—and this is the force of the priority of liberty—it is not similarly permissible to restrict political liberty in order to improve the economic conditions of the least advantaged: for example, to restrict the voting rights of the better off in order to improve the economic circumstances of the less well-off.

Second, the difference principle says that inequalities are permissible only if they maximally benefit the least advantaged. To appreciate the moral idea behind the principle, let's assume that a society guarantees equal basic liberties and fair equality of opportunity. Still, it may show considerable inequalities. In particular, suppose some people have highly marketable skills based on relatively scarce natural talents, and that others lack similarly high-end marketable skills. Assume people in both groups get up, work hard, and contribute. Nevertheless, they receive substantially different rewards in the labor market, and those differences in turn have a large impact on what they aspire to, and on the extent to which they can achieve their aspirations. How, in a society dedicated to the proposition that we all are created equal, can such inequalities, founded as they are on the contingencies of natural talent, be acceptable?

The answer provided by the difference principle is that we need to mitigate these inequalities owing to differences in natural talent. More positively stated, when the difference principle is in effect, everyone—and in particular the least advantaged group—shares in the benefits that flow from the diversity of talents in the population. "The difference principle represents, in effect, an agreement to regard the distribution of natural talents as in some respects a common asset and to share in the greater social and economic benefits made possible by the complementarities of this distribution. . . . The naturally advantaged are not to gain merely because they are more gifted, but only to cover the costs of training and education and for using their endowments in ways that help the less fortunate as well."²

The difference principle treats the distribution of talents as a common asset in that it seeks to ensure that the variety in our talents works to the benefit of all, and in particular benefit the least well-off. It does not mandate a socially beneficial use of one's talents, but does say that people can legitimately expect greater economic rewards from the use of their talents and abilities only if the use benefits the least well-off. The point of the difference principle is not to rail against the differences of natural endowments, or to eliminate them. The question of political morality is what to do with such differences, given their potentially large consequences for the fate of morally equal persons.³ The difference principle proposes an answer. In advancing the difference principle, Rawls urges, in effect, that we reject the idea that a market economy should be a kind of talent contest, designed to discover and reward the gifted. Instead, it should work as one part of a fair scheme of cooperation, and ensure a

reasonable life for all members, understood as free and equals persons: "In justice as fairness," Rawls says, "men agree to share one another's fate. In designing institutions they undertake to avail themselves of the accidents of nature and social circumstance *only when* doing so is for the common benefit."

Third, the large ambition of justice as fairness is to effect a "reconciliation of liberty and equality"—to bring elements of both liberal and egalitarian political thought together into a single coherent political philosophy. To see how, consider how the two principles work in combination. Assume first that what matters to people is not only to have legally protected liberties, but for those liberties be valuable: for them to be worth something. Assume, second, that the value of a person's liberty is importantly determined by the resources available to that person for using the liberty. In particular, assume that the worth or value of my liberties to me is an increasing function of the resources over which I exercise control: as my command of resources increases, I can do more with my liberties.

Now put the two principles together: the first ensures equal basic liberties; the second guarantees that the minimum level of resources is maximized. If the worth of a person's liberty—its value to the person—is an increasing function of the level of her resources, then by maximizing the minimum level of resources, we also maximize the minimum worth of liberty. Thus the two principles together require that society "maximize the worth to the least advantaged of the complete scheme of equal liberty shared by all." Maximizing the minimum worth of liberty "defines," Rawls says, "the end of social justice."

3. What are the arguments for the two principles?

Rawls's main argument for the two principles of justice is based on the idea of a social contract: the two principles would, he says, be agreed to in an initial contract among the members of a society about the principles to guide their society. But Rawls also offers another, more *informal* argument for the second principle. And I want first to consider that informal argument, which we find in secs. 12-13 of *Theory of Justice*.

Both arguments driven by the fundamental moral idea in justice as fairness: the idea that justice requires that we treat individuals with respect as free and equal persons, and that such treatment requires that society mitigate the effects on people's lives of the "arbitrariness of natural contingency and social fortune." The problem is to connect this abstract idea—the ideal of a fair society, whose members are treated as free and equal—with specific principles of justice. Earlier in the course, I called this "Lincoln's problem": what is it for a society to be conceived in liberty and also dedicated to the proposition that we are created equal. The contract argument and the informal argument represent two ways to bring out the content of this ideal of fair cooperation among free and equal persons: two ways to address Lincoln's problem. But they develop this idea in different ways.

In the informal argument, we start from the intuitive idea that if there are socio-economic inequalities they should be attached to positions "open to all" (equality of opportunity), and should serve the "common advantage." How should we interpret this principle? Rawls considers three conceptions of justice that offer different interpretations of the two notions in the principle: "open to all" and

"common advantage." He calls those interpretations the "system of natural liberty," "liberal equality," and "democratic equality." The central claim of the informal argument is that democratic equality provides the best interpretation of the intuitive idea. Rawls aims to bring us to the idea of democratic equality by examining certain difficulties in the other two: basically, from a very formal understanding of equal opportunity to a more demanding idea of equality. (As you will see, the three systems correspond to the ideas of the capitalist ethic, starting gate equality, and democratic equality that I alluded to at the end of the discussion of Friedman's libertarianism.)

4. Why not natural liberty (NL)?

Consider first, then, the system of natural liberty—a Friedman-esque view. Think of this view as saying the following: that a distribution of resources is just if and only if it results from a social system with: the equal basic liberties required by the first principle of justice, "Formal equality of opportunity": positions are open to all in the sense that there are no legal obstacles to attaining those positions; and an economy that operates for the common advantage (or general welfare) in that it works efficiently: no change in the institutions could make some people better off without making others worse off.

Suppose these three conditions are met. Many possible distributions of income and wealth could emerge; some will be more equal, some much less equal. According to the natural liberty conception, whichever distribution results from the choices people make under conditions that meet the conditions noted earlier is just. The fundamental idea is that if a society ensures liberties and

equal opportunity, then the distribution of resources will depend on people's choices—not artificial limits imposed by government. The distribution is just *because* it reflects these choices.

The natural liberty view rejects any regulation of choice in the name of more substantial equality of opportunity or of outcome among citizens. So inequalities may well reflect inheritance and natural talent, as well as preference and aspiration, and simple good luck. Put otherwise, under natural liberty, how people fare depends on individual choices, but those decisions are made from very different initial starting positions: different social positions, different native endowments, and different motivations that develop in light of that position and those endowments.

Why might someone object to this ideal? Recall our discussion of why we might reject a feudal system, which lacks formal equality of opportunity, in favor of natural liberty. The most plausible rationale is that opportunities should not be determined simply by social background, as they would be in a feudal system with legal obstacles on citizens in virtue of their social class. If citizens fare differently in their lives, we want those differences to have a rationale—that they not be arbitrary. And the fact that they reflect differences in social class—rather than qualifications for positions, or choices they make in light of their values—does not provide a rationale.

But this rationale implies that eliminating legal barriers to opportunity is insufficient. For a person's social background can obstruct his/her opportunities, even if positions are legally open to all. Suppose, for example, that those positions—doctor, engineer, lawyer, teacher—have high training costs. Suppose,

too, that the ability to bear those costs depends on one's social background (e.g., parental income). Then, if we hope to free life chances from class background, the system of natural liberty is inadequate. It aims only to eliminate legal barriers to opportunity, even though the reason for condemning those barriers carries over to the social obstacles as well. So it is an unstable alternative to feudalism and liberal equality.

5. Why not Liberal Equality (LE)?

To mitigate this dependence of life chances on social background, Liberal Equality adds a condition of *fair equality of opportunity*: *People who are equally endowed and motivated ought to have equal chances of attaining desirable positions and offices*. In short, “The expectations of those with the same abilities and aspirations should not be affected by their social class” (63).

We get to liberal equality by shifting the interpretation of the idea that positions must be “open to all” from formal equality of opportunity to the more substantive Fair Equality condition. According to liberal equality, then, a distribution is just if and only if it results from choices in a social system that meets four conditions: equal basic liberties; formal equality of opportunity; fair equality of opportunity; the economy operates for the common advantage, in the same sense as earlier. So when the liberties and fair equality are ensured, then any distribution that emerges reflects the choices people make and so is a just distribution, regardless of the dispersion in the distribution. We have a strong equal opportunity requirement, but no direct limits on inequalities of outcome.

But liberal equality itself may still be troubling, and in two ways. First, assume that your social background shapes your aspirations, motivations, and realized abilities: Rawls says, not implausibly, that "Even the willingness to make an effort, to try, and so to be deserving in the ordinary sense is itself dependent upon happy family and social circumstances" (74). Of course the dependence here is not deterministic, but it still may be real. But if that is true we will be dissatisfied with the system of liberal equality. Why?

Liberal equality seeks to prevent social background from determining life chances: that's why it is preferable to natural liberty. But it allows that life chances may reflect aspirations, motivations, and realized abilities. But suppose our aspirations and abilities—how much effort we put in—are influenced by social background. If so, then Fair Equality itself permits social background indirectly to determine life chances. While people with equal abilities and motivations have equal chances, people born into different social classes develop different abilities and make different efforts. Then the fact that there is fair equality does not end the dependence of social advantage on social background. It simply changes the mechanism through which social background shapes life chances.

Second, under liberal equality "social inequalities exactly express natural inequalities."⁴ But if social background should not fix life chances, then why is it permissible for them to be determined by differences in natural endowment or fortune? Here we have Friedman's objection to starting gate equality: If we are troubled by the fact that differences in life chances reflect differences of social background, then we also have reason for being troubled if they are determined by differences in natural endowment. "There is no more reason to permit the

distribution of income and wealth to be settled by the distribution of natural assets than by historical and social fortune” (74).

6. Why Democratic Equality (DE)?

Like Friedman, then, Rawls points to the intellectual instability in liberal equality. But Rawls proposes to resolve that instability by defending a more egalitarian conception of justice, which he calls *democratic equality*. The central idea is to sever life chances—in particular, our position in the distribution of income and wealth—from social starting position, natural endowments, and luck. The idea is to carry through on the idea of making outcomes depend on the decisions people make about to do with their lives, rather than the resources or talents that we happen to be born with—by establishing greater equality of circumstance.

The democratic conception results from adding the difference principle to equal basic liberties, efficiency, and fair equality of opportunity. Thus a distribution is just if and only if it results from a system with basic liberties and fair equality of opportunity and in which inequalities can reasonably be expected to contribute to the maximum well-being of the least well-off. The difference principle does not limit inequalities by placing any direct restrictions on income dispersion: on the distance between top and bottom, for example. Instead, it requires that improvements at the top be tied to improvements elsewhere—in particular, to the maximum improvements at the bottom.

How might we get to this seemingly extreme requirement—extreme in attaching so much weight to the size of the minimum, and none at all to the central tendency or dispersion? Start with the idea that the distribution of

economic advantage is not a product of nature, or simply of the separate choices of individuals, but at least in part a result of the rules of the economic game that we collectively choose to adopt and enforce: a product in part of what Rawls calls the “basic structure of society.” Suppose, too that we think that those rules—which we collectively choose to adopt and enforce—ought to be founded on the idea that people are to be treated as equals, irrespective of the reflect the decisions that people make about their lives, and ought to be severed from the contingent differences that distinguish among equal moral persons. But to see us as free and equal persons is to see us in abstraction from all the features that distinguish us—all the contingencies of social fortune and luck in the natural lottery. This naturally suggests an equal distribution of income and wealth. For if we regard all the sources of differences in fate as morally arbitrary, and then strive to eliminate the effects of the morally arbitrary on the distribution of resources, the result is equality of outcome, and not simply equality of opportunity.

But now suppose an unequal distribution—by providing incentives or compensation—would increase the size of the pie, and so could improve the circumstances of all. So the inequalities could work to the advantage of all: permitting such inequalities, then, can itself be interpreted as a consequence of an equal concern for the good of each. This brings us to the second stage: what sorts of departures from equality are consistent with the idea of separating life chances from the contingencies of social and natural fortune. How are the gains that result when we move away from equality to be distributed?

According to the difference principle, the justification for some positions being more favored than others —e.g., for giving rewards to those whose natural abilities enable them to contribute more—is that by favoring them, benefits to those in less advantaged positions will be maximized. This is a way to gain the advantages to all that can result from inequalities while mitigating the effects on people's lives of the contingencies of social background, natural endowment, and good luck. It mitigates in precisely the following way: that no one is permitted to take advantage of his/her social position or natural endowments except and insofar as their doing so improves the conditions of the least advantaged. So the fact that someone is in the least advantaged group—because of the social background, or native endowments—has less effect on his/her material conditions than under alternative distributive principles.

But while the informal argument has some force, it is not clear why the two principles, and maximizing the minimum in particular, are the best way to express the idea of treating people as free and equal moral persons. Offhand, an equal distribution also satisfies that condition, as would a distribution that ensures a decent minimum without maximizing the minimum. In short, it is not clear from the informal argument that there are any determinate principles that express the ideal of a distribution of rights, duties, and advantages that nullifies or mitigates the contingencies of natural fortune, social circumstance, and simple good luck. This, then, is the role of the more formal argument for principles from the original position.

7. Why a social contract?

To defend his principles of justice, Rawls revives the social contract idea associated with Locke, Rousseau, and Kant. The social contract tradition argues that the right way to order a society is the ordering that the members of the society would *unanimously agree to*. Because of the requirement of unanimity, each person has veto power over the terms of the agreement, so the terms of the agreement must be justified to each person, who must find the basics of the society acceptable. We are to live together on terms that each of us finds acceptable. And as Hobbes said: “that which every man will have so, no man can say is unjust.”

This idea of a unanimous initial agreement on the basic terms of social cooperation has an obvious attraction for a society of equals. But at the same time, a basic fact of social life is that people disagree about issues of morality, politics, and religion. How could there ever be unanimous agreement? If there is to be agreement, we need to impose some special conditions on the agreement. We cannot simply take a vote: that will simply reveal the points of disagreement, not generate a unanimous agreement. At the same time, not just any conditions designed to generate agreement will do. We might be able to get a unanimous agreement if we injected everyone with a drug that induces head-nodding and thus secures agreement, but that would not justify the results.

How then can we navigate between voting and drugging? One thought is that people who disagree about what justice demands nevertheless agree, or might be brought on reflection to agree, on certain fundamentals. Suppose, then, that we could use these fundamental points of agreement to define the

circumstances in which people make a social contract—to set acceptable conditions on the circumstances of agreement. Then perhaps we could get unanimous agreement about basic principles of justice. So what might the points of agreement be, and how could those be expressed in the circumstances of agreement?

Rawls suggests three points of agreement:

1. First, that certain particular practices are unjust—e.g., religious intolerance and racial discrimination. Convictions about the injustice of these practices are, Rawls says, "provisional fixed points which we presume any conception of justice must fit." So the social contract condition is designed so that the parties will agree on these points, whatever else they agree on. It is simply built into the situation that the result has to fit there "data points."

2. Second, about a "conception of the person." By a "conception of the person," Rawls means a view about the features of human beings that are of fundamental importance and relevance about when it comes to addressing questions of public justice. For the utilitarian, what is most fundamental is our capacity to experience pleasure and pain. For justice as fairness, what matters most when it comes to issues of justice are not the differences among us—differences of race and sex, social background, native talents, and religious, moral, and personal ideals. These differences, as important as they are in some settings, we should regard, Rawls says, as "accidents of natural endowment and contingencies of social circumstance;" such accidents are "arbitrary from a moral point of view." Instead, what is relevant is that we are free and equal moral persons. Let me explain this important idea:

First, individuals are assumed to have a conception of the good—a set of goals, attachments, and loyalties, which serve to guide their conduct.

Conceptions of the good may be more individualistic or more communal: the essential point is that individuals have them and they vary across people.

Second, we have the capacity to form a conception of the good—that is, a capacity to decide on, to pursue, and to revise our conception of the good. We might undertake such revision in the light on new information, wider experience, new forms of self-understanding, and moral or religious reflection. We are neither unconditionally committed to our current view of the good, nor are we simply a bundle of unrelated preferences and goals; instead we have aims and aspirations, and are able to revise them on reflection.

Third, we have the capacity for a sense of justice, i.e. to grasp the principles specifying fair terms of social interaction and to guide our conduct in light of that understanding.

These common potentialities—"moral powers," in Rawls's term—define us as free and equal moral persons. We are *equal* in that each has, to a sufficient degree, these basic powers, which enable us to participate as full members of the society. This is not of course to deny our many differences: of course we are very different from one another as well. But the point is to affirm that those differences—of talent and ambition, religious and gender, race and ethnicity—do not touch our standing as equals; for that status, the possession of the moral powers is sufficient. Moreover, as possessors of the basic moral powers we are *free*. In particular, we have and are recognized as having the capacity to alter our goals, attachments, and loyalties without losing standing as citizens—for

example, to undergo religious conversion or change of political commitment without loss of rights.

3. Finally, there is agreement about an abstract ideal of social cooperation—in particular, about the importance of fairness in society. While people have different ends and goals, different backgrounds and talents, we each ought to have a fair chance to develop our talents and to pursue those goals.

So here is the thought: despite disagreeing about what justice requires, we might be brought on reflection to agree, as a basis of political argument, on these basics: people who argue about what justice requires might be thought of as arguing about what are fair terms of cooperation between free and equal persons. Rawls's idea is to take these points of agreement, and construct a conception of justice around them by building the social contract to reflect them. Though we now disagree, perhaps deeply, about what a just society is, we agree—or could be brought to agree on reflection—with the very abstract idea that justice requires a society that is fair to its members considered as free and equal moral persons, a society whose basic structure works to "nullify the accidents of natural endowment and contingencies of social circumstance as counters in [the] quest for political and economic advantage." The problem is to ensure that the initial contract reflects this ideal.

8. What is the Original Position (OP)?

So the task is to connect the abstract ideal of fair cooperation among free and equal persons—each with the basic powers required for full participation in

society—to specific requirements of justice. And Rawls proposes to bridge this gap through the social contract: in particular, a contract built around the points of agreement.

Rawls asks us to imagine a hypothetical situation—the Original Position—in which people are to choose principles of justice for their own society. That contract situation is constructed to reflect the fundamental points of agreement, in particular, the conception of the persons as free and equal. The essential idea is that certain of our characteristics are not relevant in deciding what we are entitled to as a matter of justice. To represent that idea of irrelevance, Rawls proposes that we make the choice of principles behind a "veil of ignorance" in which we are *assumed to be unaware of* the irrelevant features. Behind the veil, we do not know, for example, whether or not we are blessed by natural chance, or whether the contingencies of social circumstance are favorable or unfavorable.

The veil of ignorance can seem like a very artificial device, and in a sense it is: how could we not know our social position, our gender, our values and religious convictions. But keep the point in mind: this is an analytical model, designed to express a moral idea, and to make our reasoning about justice more tractable. The moral idea is that certain factors are irrelevant to justice. The veil of ignorance proposes that we *model irrelevance by ignorance*. In particular, we do a thought experiment. Imagine, hypothetically, that we are to make a choice of principles of justice for our society on the assumption that we, as the parties making the choice, do not know our social class position, natural talents, sex or race, conception of the good, or anything else that distinguishes any one of us from other free and equal moral persons. Because the parties must reason as if

they did not know these things, they will not be able to tailor principles to advantage themselves, or members of their class, sex, race, religion, or moral tradition. Not knowing these facts, people have a chance of arriving at a unanimous agreement. Though conditions like the veil of ignorance may seem artificial, the idea is simply “to make vivid to ourselves that it seems reasonable to impose on arguments for principles of justice, and therefore on those principles themselves” (16).

Rawls proposes, then, that if people reason about issues of justice as though they were unaware of the social contingencies and the accidents of natural endowment, then they would choose his two principles of justice—with their assurance of maximin worth of liberty—over alternative conceptions.

Because the aim of the original position is not simply to reach an agreement (we can get that with a head-nodding drug), but to find principles suited to the ideal of fair cooperation among free and equal persons, we are to place behind the veil of ignorance all the features that distinguish among free and equal moral persons—including their religious ideas, moral philosophies, and views of social justice. These ideas are very important to us: to who we are, and how we conduct our lives. But they distinguish people, understood as free and equal citizens, so we put them aside. The parties in the original position know only that they represent the interests of a person who has some conception of the good, perhaps a religious outlook (though they do not now what that conception is); a person who has an interest in being able to choose and revise their ends; and who has an interest in forming and acting on a sense of justice.

Once they know that, however, they also know that advancing those basic interests requires certain goods—"social primary goods"—and so the parties to the social contract know that they need these goods. In particular, the social primary goods are:

- i. The basic liberties, including freedom of thought and conscience, the political liberties, liberty of association, the liberties associated with the integrity of the person, and the liberties associated with the rule of law.
- ii. Freedom of movement and choice of occupation, under conditions in which there are a variety of opportunities.
- iii. Powers and prerogatives of office and positions of responsibility.
- iv. Income and wealth.
- v. Social bases of self-respect.

What is special about these goods? Given the conditions of social cooperation among human beings, free and equal citizens *need* these goods whatever their particular conception of the good may be. They need them because these goods are required for pursuing a wide range of ends, and for developing and exercising the potentialities or basic "moral powers" that define a moral person. Of course, we need other goods as well, but these social primary goods are more directly dependent on social institutions than are other primary goods ("health and vigor, intelligence and imagination").

For example, consider the basic liberties. Liberty of thought and conscience is a primary good in part because it is required for the pursuit of the various moral, religious, and philosophical conceptions that serve to support our conceptions of the good: in particular, it is required if we are to fulfill the

obligations that our moral and religious views (if we have them) assign to us. Or in the case of income and wealth: we typically need resources to pursue our aims and ambitions. The case of self-respect is especially important, and I will come back to it later on.

10. What is the intuitive idea behind the OP argument?

Under the veil of ignorance, the parties base their choice of principles on the consequences of those choices for their level of primary goods. The central claim, then, is that parties in the original position would prefer their expected level of primary goods under the two principles than under any of the alternatives, and therefore would choose those principles.

The argument itself is complicated, so it will help to have the intuitive line of thought in mind: You are asked to choose principles that will regulate the society you live in, and you understand that you will not be able to change your mind after the fact. You will make the choice under conditions of ignorance about yourself, your ideals, and your social position. Because you do not know which person you will be, but have to live with the principles you choose, you want to be sure—if this is possible—that your situation is (roughly) acceptable whatever it turns out to be. Because of the veil of ignorance—remember, it models moral equality as ignorance of who you are—you want to be sure that the society is acceptable from the point of view of each person, because you may be that person. In particular, you want to be sure that it will be acceptable even if you land in the lowest social position, where it is least likely to be acceptable. And, according to Rawls, this is just the insurance—the strong downside protection—

that the two principles provide: they ensure that social arrangements are acceptable to all members of a society of equals.

But why focus so much on downside protection? To see the force of the question, consider the contrast between Rawls' two principles and the principle of average utility. According to the principle of average utility, an action or institution is right or ought to obtain just in case it maximizes the sum of utility divided by the number of people, rather than just the sum of utility (with evident differences for questions of population policy).

A society regulated by either justice as fairness with its two principles or by the principle of average utility would include a range of different levels of primary goods, associated with different social positions. Let's focus for a moment only on the economic implications of the principles. Thus, the minimum income under the two principles is a maximized minimum; it must be at least as high as the minimum level in a society that aims not to maximize the minimum but to maximize average utility (assuming that utility is increasing in income). But because the average under AU is a maximized average, it must be at least as high as the average under the two principles; moreover, the maximum level may well be higher. That is, the average level of utility may well be greater if a society permits greater dispersion of circumstances and allows inequalities that do not contribute to the well-being of the least well-off group.⁵

Here, then, is the question: Suppose you are reasoning from behind the VI: you do not know where you will end up, but you do know that there are a range of possibilities. Is it rational to take a chance with maximizing the average—with focusing on the central tendency—knowing that the worst

situation will almost certainly be less good, or to opt for the two principles? Notice two considerations that are important to the decision, and they work in different directions. If you opt for justice as fairness, you buy strong protection against downside risks: in effect you buy insurance against luck, or inheritance, or talent not working out well, since you ensure that the minimum is as high as possible. But insurance has a cost: if things work out well, you may end up doing less well than you might otherwise have done—you pay the premium but do not collect. How then are parties in the OP to balance the high security level provided by the two principles against the possible gains that could be won by choosing average utility?

11. How does self-respect lead to the two principles?

So what leads the parties in the original position to be so concerned about ensuring that the worst situation is as good as possible? The rationale has a great deal to do with the stakes of the choice, which are supposed to make us especially concerned about the worst eventualities, and assuring that those worst eventualities are as good as they can be. And one consideration that brings those stakes out with particular force is Rawls's account of self-respect: the concern with self-respect provides a way to see how a range of different considerations come together in favor of justice as fairness. The point, in essence, is that when we make our choice in the original position, we need to be very conservative in our choice—it makes sense to avoid risks and focus focus on the worst situations—in part because a profoundly important good is at stake in our decision, namely our self-respect. In a nutshell, our chances of living a

good life will depend on achieving self-respect; but our achieving self-respect will depend in part on how we are regarded and treated by others in our society; how we are regarded and treated by others in our society is expressed in the public understanding of justice for the society, and justice as fairness provides an account of justice that provides strong support for our sense of self-worth, no matter where we end up.

What is self-respect? According to Rawls, self-respect has two main elements: (i) that a person has aims and aspirations—a conception of the good—and has reasonable hopes for success in achieving those aims; (ii) that a person has a sense of his own worth, and an associated conviction that his/her aims are worth achieving (386). But why more precisely is self-respect relevant to the choice of principles? That relevance reflects three features of self-respect.

The first is that self-respect is a fundamental good—perhaps the fundamental primary good—whose presence is almost certainly required if conditions are to be acceptable. Without self-respect, "All desire and activity becomes empty and vain, and we sink into apathy and cynicism" (386). Or again: "A sense of their own worth is necessary if they are to pursue their conception of the good with satisfaction and take pleasure in its fulfillment" (155). If self-respect is a fundamental good in these ways, then it will be of particular importance for persons making the choice of principles in the original position: "they would wish to avoid *at almost any cost* the social conditions that undermine self-respect" (emphasis added). If it really makes sense to avoid "at almost any cost" the conditions that undermine self-respect, then you need to make a choice in the OP that does not jeopardize self-respect. And this importance suggests the

following test that parties might use in assessing principles in the original position: consider whether each position you might end up in under the principles you are assessing ensures conditions favorable to the self-respect of the person in it.

But even if we assume that self-respect is of profound importance in these ways, it may nonetheless seem strange to treat self-respect as a social primary good, and to suppose that it is relevant to the original position. After all, self-respect is not something that society has to distribute to people: it is not like income, or wealth, or liberties, or opportunities, dependent in some straightforward way on social and political institutions. But—here I come to the second main idea—while self-respect itself is not socially distributed, certain social conditions that support self-respect are: Rawls calls these conditions the "social bases" of self-respect, by which he means social conditions that foster self-respect.

It is useful to distinguish two kinds of social bases, corresponding to the two aspects of self-respect I distinguished earlier: (i) It is important that our circumstances enable us both to develop aspirations—a conception of the good—and to pursue that conception with reasonable prospects of success: call this the "objective social bases" of self-respect; (ii) It is important that we associate with others in such a way that we experience their respect for us; "self-respect normally depends on the respect of others" (155), and without such experience, our sense of our own worth and that our aspirations are worth pursuing may be hard to sustain. Call this the "interpersonal social bases" of self-respect.

Third, the other primary goods (other than self-respect)—basic liberties, opportunities, income and wealth, powers and positions of authority—provide the social bases of self-respect, both objective and interpersonal. Later I will explain in more detail how this connection between primary goods and social bases works. Suffice it to say now that the concern with the fundamental good of self-respect, so important to leading a good life, leads to a concern with the social bases of self-respect, as supporting conditions for self-respect. And the concern with social bases, both objective and cultural, leads to a concern with the distribution of the other primary goods since they provide those social bases. So if we need to be sure, so far as possible, that we make a decision that does not risk self-respect, then we need to be especially concerned with the worst eventualities—the lowest levels of primary goods—because that is where self-respect is most likely to be threatened.

Self-respect has very great value, then, and it has social foundations, of these two kinds. So the parties in the original position cannot make a choice that jeopardizes their self-respect. To ensure that they have an acceptable situation, then, the parties in the original position need to ensure that social conditions will be supportive of their self-respect. Consider, then, how a concern for the social foundations of self-respect leads to a case for the two principles.

Thus, it counts in favor of the two principles that they provide equal basic liberties, basic liberties at every social position including the minimum. Why? First, because having a right to the liberties is itself an expression of respect for each person: they are entitled to develop and pursue aspirations, whatever position they are in, and without needing public approval of others for the value of

what they are doing. Moreover, second, because the liberties enable citizens, whatever their social position, to form aspirations and to associate with others in pursuit of common ideals: such common pursuits in association with others who share their concerns and ideals will help to provide what I have been calling the interpersonal bases for their self-respect. By ensuring the liberties to all, then, regardless of their social position, the two principles help to ensure in particular that the minimum position is satisfactory or acceptable because they help to establish the conditions of self-respect at that position. In contrast, utilitarianism does not provide comparable assurance of basic liberties, and thus is a much more risky choice.

Considerations of self-respect also give the parties reason for concern about the material resources available to them. They want to ensure that they can pursue their aims with self-confidence: the objective basis of self-respect. But for that purpose they need more than simply the liberties to pursue those aims. They need to be sure that their liberties are worth something, that the liberties are not just formally protected, but also genuinely valuable. So they want to ensure what Rawls calls the "worth of liberty" (179): a decent standard of living that will enable them to use their liberties in pursuit of their aspirations, no matter what position they end up in.

But concern about a decent standard of living—say, a reasonable floor under income, or a threshold—is one thing; ensuring that the minimum is maximized is another. Why do considerations of self-respect lead to the difference principle as a way to ensure that decent standard of living?

Consider how the difference principle provides particularly strong support for self-respect at the least advantaged position.

1. The difference principle only permits inequalities that contribute to lifetime expectations at the least well-off position. Smaller inequalities would reduce expectations, as would greater inequalities.

2. Assume now that the value or "worth" of liberties to a person depends on the level of resources available to the person. (This is much less plausible for the case of the political liberties. But their worth is ensured by the proviso in the first principle requiring the fair value of political liberty.)

3. Therefore, the minimum value of the liberties with the difference principle is greater than the minimum value under the alternatives. Indeed, if the value of the liberties is, as proposed, an increasing function of (absolute, not relative) income and wealth, then the value of the liberties achieves its maximum when the difference principle is satisfied.

4. But self-respect depends on the value of the liberties. For self-respect requires confidence in one's ability to successfully pursue one's aims: it requires objective supports. But that confidence is increased as resources increase. In short, then, the great value of self-respect encourages the choice of principles that maximize the minimum worth of liberty.

5. So the two principles together provide strong insurance for the social bases of self-respect, even at the minimum position.

Furthermore, under the difference principle occupants of the least well-off position are not only guaranteed the worth of the liberties; the principles of justice are part of the public culture in a just society, so this guarantee of the worth of

liberties is part of a public understanding in the society. So the society of justice as fairness protects liberties and in effect agrees to ensure advantages regardless of the particulars of social position, natural endowment, and good fortune that distinguish the free and equal members of a well-ordered society. To forgo possible advantages because accepting them would reduce expectations at the minimum, and the worth of liberties at the minimum, is to express respect for those at the minimum position and fully affirms their worth. And in view of the importance of self-respect, and the interpersonal bases of self-respect, it is rational to want public institutions to show respect for people. “Thus a desirable feature of a conception of justice is that it should publicly express men’s respect for one another. In this way they ensure a sense of their own value.” Now “by arranging inequalities for reciprocal advantage and by abstaining from the exploitation of the contingencies of nature and social circumstance within a framework of equal liberties, persons express their respect for one another in the very constitution of society” (156).

So the choice of the difference principle, alongside the requirements of liberty and fair equality, strengthens the foundations of self-respect both by ensuring the resources required for the self-confident pursuit of aims (the objective aspect of the social bases of self-respect) and by contributing to the experience of respect (the interpersonal aspect of the social bases of self-respect). It builds a respect for each member of the society, as free and equal, into the basic principles regulating the society.

In effect, what Rawls has said in this argument is the following: in a modern political society, the distribution of income is not natural and fixed, or

simply the result of separate individual decisions. Instead, it results in part from our collective decisions, about the rules of the economic game concerning property, contract, taxation, education, training—our collective decisions about laws and policies. Because outcomes are dependent on the rules we adopt collectively—on the basic structure of the society—and not simply the choices we make individually, the stakes in ensuring a fair distribution of income are very high: in particular, increasing material well-being for citizens who are least well-off is not simply a matter of increasing their welfare. Instead, it is a way of providing a kind of social recognition of their equal worth, by ensuring them the resources they need to formulate and to pursue their plans of life on an equal footing.

So just as we show respect for members as equals by providing rights to personal and political liberties—an entitlement to pursue personal aims and to participate as equal members of the sovereign people in deciding the rules of the society—so too we show respect for members as equals by establishing rules of the economic game that foster a fair distribution of the resources people need to pursue their aspirations, in particular by ensuring that we maximize the value of liberty to those for whom the value is lowest. Under any structure of laws we can imagine, the prospects of some citizens will be less attractive for their entire lives than the prospects of other citizens. In a society of equals, we need to be able to explain, to those whose prospects are worse, why we have not chosen a different arrangement under which their prospects would be better. Under Rawls principles, the explanation is that the prospects for those who are worst off are as good as they can be.

12. Conclusions

Rawls's basic contention is this: Start from the fundamental ideal of a fair system of cooperation among free and equal moral persons, who have divergent moral and religious convictions. Consider which principles of justice are suited to fair cooperation among persons thus understood. Beginning with that basic ideal, we will want to find principles acceptable to each person, whatever his/her social position, talents, sex, race, religious creed, aspirations. Acceptability turns crucially on support for self-respect. And if we want to find such principles, then we will arrive at the two principles, those principles having the property that would be acceptable even if we turned out to be the least advantaged.

Starting from the fundamental ideal of fair cooperation among free and equal moral persons, we will be led to a political conception mandating both security of basic liberties and socio-economic equality: a "reconciliation of liberty and equality." What this suggests is that we ought not to think of liberty and equality as independent values that need to be balanced much less as warring principles, nor should we suppose that the combination of the two into a single conception of justice is simply a matter of political compromise. Instead, liberty and equality have a common foundation in the ideal of a society that recognizes its members as free and equal, and provides fair conditions of cooperation among them. That ideal both supports such fundamental convictions as that religious intolerance and racial discrimination are unjust and at the same time provides egalitarian guidance on controversial questions about the distribution of social and economic resources.

¹ *TJ*, p. 86.

² *TJ*, 87.

³ “The natural distribution is neither just nor unjust; nor is it unjust that persons are born into society at some particular position. These are simply natural facts. What is just and unjust is the way that institutions deal with these facts. Aristocratic and caste societies are unjust because they make these contingencies the ascriptive basis for belonging to more or less enclosed and privileged social classes. The basic structure of these societies incorporates the arbitrariness found in nature. But there is no necessity for men to resign themselves to these contingencies” (87-88).

⁴ *Ibid.*, p. 377.

⁵ This is a rough approximation to the difference: it needs to be adjusted for the fact that the two principles specify a distribution of primary goods, while average utility operates in utility space.