

Philosophy 285 Fall, 2007 Dick Arneson
 Overview of John Rawls, *A Theory of Justice*

Views of Rawls's achievement:

G. A. Cohen: "I believe that at most two books in the history of Western political philosophy have a claim to be regarded as greater than *A Theory of Justice*: Plato's *Republic* and Hobbes's *Leviathan*. I shall not try to say what I think is great about those books. But among what contributes to the greatness of *A Theory of Justice*, and of the entire Rawlsian achievement, is that, to put the matter as Hegel would have done had he agreed with me, John Rawls grasped his age or, more precisely, one large reality of his age, in thought. In his work the politics of liberal (in the American sense) democracy and social (in the European sense) democracy rises to consciousness of itself."

Brian Barry: "Nothing Rawls has written since about 1975 is any good." {I can't find the reference, or the exact quote, but I'm sure he wrote roughly this somewhere.]

1. There are two big ideas in Rawls: (1) justice is the outcome of a suitably defined social contract. What justifies candidate principles of justice (fundamental principles of morality) is that they would be agreed to in the "original position," Rawls's term for the philosophically most favored interpretation of the original position. (2) A particular interpretation of liberalism as the substance or content of justice. Liberalism is the conjunction of two ideals—liberty and democratic equality. Each of us has her own life to lead, and should be left free to govern her own life—hence freedom of thought and religion, freedom of speech, wide personal liberty of action. The other ideal is that of a society of equals, a democratically self-governing community of people undivided by distinctions of caste status and unmarred by excessive inequality of life prospects. Rawls's two principles of justice interpret and specify these ideals.

2. The enterprise of developing and defending his conception of justice takes the shape it does in *A Theory of Justice* by being organized by Rawls's conviction that the chief rival, the theory to beat, is utilitarianism. This doctrine holds "that society is rightly ordered, and therefore just, when its major institutions are arranged so as to achieve the greatest net balance of satisfaction summed over all the individuals belonging to it." (22). In contrast, Rawls holds "Each person possesses an inviolability founded on justice that even the welfare of society as a whole cannot override" (3). In a gnomic phrase, Rawls observes, "Utilitarianism does not take seriously the distinction between persons." What does this mean?

Part of the attraction of utilitarianism according to Rawls is that its principles define a genuine theory—the principles pick out the facts (information) relevant to choice of policy and act, and given a full specification of these facts, the principles deductively determine what is to be done. Rawls sees the prevailing alternative to utilitarianism as the no-theory-theory, what he calls intuitionism, as in W. D. Ross. Morality identifies plural considerations relevant to choice of action and policy, but these considerations do not by themselves determine what is to be done even when all conceivably relevant facts are specified. Rawls thinks you can't beat a theory except with another, better theory. Rawls aims to develop a genuine theory fit to supplant the utilitarian theory.

Part of the attraction of utilitarianism according to Rawls is its teleological character. In a teleology, the right is defined independently of the good, and the right is then identified with maximizing the good. This view fits with the intuitive notion that rationality is maximizing. The rational agent maximizes the satisfaction of her goals. The question then becomes, what goals are rational to pursue. If one supposes rationality requires that one's own interests be the goal, then we get prudential rationality. The prudent agent acts so as to maximize the satisfaction of her own interests. Morality is identified with the thought that the goals one should pursue are fixed from an impartial perspective. From this perspective, each person's interests count the same as the comparable interests of anyone else; each person counts for one and nobody for

more than one. The moral goal is then the satisfaction of the interests of all (people or other sentient creatures). Rationality as maximizing plus morality as adoption of the impartial perspective then yields utilitarian teleology. Rawls proposes to appeal to the Kantian tradition to turn this around: in a slogan, the right is prior to the good. (Either the good is not defined independently of the right, or the right is not identified with maximizing the good, or both.)

3. The primary subject of justice, according to Rawls, is the “basic structure” of society—“the way in which the major social institutions distribute fundamental rights and duties and determine the division of advantages from social cooperation.” Why focus on the basic structure? Here’s what Rawls says. “The basic structure is the primary subject of justice because its effects are so profound and present from the start. The intuitive notion here is that this structure contains various social positions and that men born into different positions have different expectations of life determined, in part, by the political system as well as by economic and social circumstances. In this way the institutions of society favor certain starting places over others. These are especially deep inequalities. Not only are they pervasive, but they affect men’s initial chances in life; yet they cannot possibly be justified by an appeal to the notions of merit or desert. It is these inequalities, presumably inevitable in the basic structure of any society, to which the principles of social justice must in the first instance apply” (section #2, p. 7).

There is a contrast here between deep and shallow inequalities. The latter are mediated by choice, for which one may be responsible in whole or in part. The former are thrust on people independently of their choices. The operation of the basic structure generates deep inequalities, e.g. between someone born in wealth or poverty, or one born a Brahmin or an Untouchable. Notice that there seem to be deep inequalities that are generated at least largely independent of the basic social structure—inequalities in different people’s genetic endowments and hence their differential in-born proclivities to develop talents and virtues, and also differences in the quality of the early socialization one receives. These forces would operate even in a state of nature.

One might see moral obligations including justice obligations as essentially obligations of individuals to act in one or another way. Institutions in this perspective are seen as devices that we perhaps should build in order to facilitate the fulfillment by each of us of our individual conduct obligations. Rawls’s idea is different—the principles of social justice are in the first instance principles for the regulation of the basic structure.

4. In chapter 2 of *A Theory of Justice* Rawls asserts and explicates his two principles of justice and offers intuitive rationales. These intuitive justifications, he says, are provisional; the heavy lifting of justification of principles is supposed to be done in the original position arguments of chapter 3. But the chapter two justifying remarks are interesting, and will especially interest those of us who find the original position arguments less than compelling.

The two principles: The Equal Liberty Principle: “each person is to have an equal right to the most extensive basic liberty compatible with a similar liberty for others” (p. 60). The basic liberties are given by a list: “The basic liberties of citizens are, roughly speaking, political liberty (the right to vote and to be eligible for public office) together with freedom of speech and assembly; liberty of conscience and freedom of thought; freedom of the person along with the right to hold (personal) property; and freedom from arbitrary arrest and seizure as defined by the concept of the rule of law.”

The second principle: “Social and economic inequalities are to be arranged so that they are both (a) to the greatest benefit of the least advantaged and (b) open to all under conditions of fair equality of opportunity” (p. 83; section #13).

The first principle has strict lexical priority over the second and part b (fair equality of opportunity) of the second principle has strict lexical priority over part a (the difference principle) of the second principle. Lexical priority of A over B means one should do whatever one can to attain A to the maximal degree possible and turn to effort to fulfill B only when and to the extent doing so does

not diminish even to the slightest degree the extent to which A is fulfilled. So, no trade-offs; no acceptance even of the slightest violation of the basic liberties to gain even the greatest boost in the extent to which fair equality of opportunity, or as a still lesser priority the difference principle, is fulfilled. For more on the priority relations, see chapter 4, section #39. For a justification of the priority of basic liberty, see section #82 in chapter 9. So far as I can see, the book contains virtually nothing by way of justification of fair equality of opportunity (but see the brief remarks on page 84).

Fair equality of opportunity obtains in a society when any two persons with the same native talent potential and the same ambition have the same prospects of success in fulfilling those ambitions (to seek positions that confer extra gains of basic resources of primary social goods such as income and wealth). Rawls writes, elucidating the ideal, "assuming there is a distribution of natural assets, those who are at the same level of talent and ability, and have the same willingness to use them, should have the same prospects of success regardless of their initial place in the social system, that is, irrespective of the income class into which they are born" (p. 73). Despite the last qualifying phrase, I suppose Rawls means that nothing—not one's racial identification, ethnicity, sex, income and wealth of parents, and so on—except one's native talent and one's ambition should be predictive of social success. Any two people with the same ambition and same native talent have the same prospects of success. This is the ideal of a classless society in a sense. From one's initial placement into the social system, nothing can be predicted about one's future. Being born on the right or the wrong side of the tracks makes no difference to one's prospects

Qualification: Fair equality of opportunity might be limited to an extent by the prior basic liberty principle, which presumably entails that each individual is free to choose her own mates and romantic and marital partners. If society could better fulfill fair equality of opportunity by a social assignment of marriage partners, the equal liberty principle rules that out. But given basic liberty to cohabit and marry, won't parents inevitably successfully seek to give their own children a leg up in social competition? When fair equality of opportunity obtains, parents can do this, but their efforts are entirely offset by other social circumstances. (For example, society puts in place special education programs for children from disadvantaged families that succeed in completely offsetting the disadvantages. Also, fair equality of opportunity might conceivably set limits to allowable inequality of income and wealth. Perhaps if the gap between the income of the highest and lowest decile of the population rises too far, fair equality of opportunity could not be fulfilled—then Rawls would be committed to capping allowable inequality of income.

We can contrast formal equality of opportunity (careers open to talents) and formal equality plus. Formal equality of opportunity obtains when desirable jobs, entrepreneurial opportunities, bank loans, places in universities and colleges and universities, and so on are open for anyone to apply, with applicants being assessed by their merits and the relevantly most meritorious candidate being selected. Formal equality of opportunity could obtain in a society in which there are two classes, one wealthy, one not, such that only members of the wealthy class have the opportunity to become qualified for desirable positions and do become so qualified. Formal equality plus requires that something be done to move toward equalizing people's opportunities to become qualified. Free public education financed by general taxation is a step in the direction of formal equality of opportunity plus. Rawlsian fair equality of opportunity takes the idea of formal equality plus to its logical limit.

5. The difference principle. Behind equal basic liberty and fair equality of opportunity in Rawls's nested set of priorities is the difference principle. This says social and economic inequalities are to be set so they work to the maximal advantage of the least advantaged group. (Qualification: there is also the shadow just savings principles that Rawls sees as constraining the degree to which justice requires shifting resources now to the worst off. See chapter 5.)

Why does justice according to Rawls demand the difference principle? In section #12 of chapter 2 he considers an alternative, a competitive market economy operating under the standard

textbook conditions, so the result is efficient: no one can be made better off without making someone else worse off. Assume equal liberty is in place, and at least formal equality of opportunity (careers open to talents). Let each make deals as she chooses and cooperate with others on mutually agreeable terms, in an environment where all have the same freedom. Perhaps we should regard the outcomes of such a competitive market economy as fair, whatever they turn out to be.

Rawls observes of this arrangement, which he calls the system of natural liberty: But since there is no effort to preserve an equality, or similarity, of social conditions, except insofar as this is necessary to preserve the requisite background institutions, the initial distribution of assets for any period of time is strongly influenced by natural and social contingencies. The existing distribution of income and wealth, say, is the cumulative effect of prior distributions of natural assets—that is, natural talents and abilities—as these have been developed or left unrealized, and their use favored or disfavored over time by social circumstances and such chance contingencies as accident and good fortune. Intuitively, the most obvious injustice of the system of natural liberty is that it permits distributive shares to be improperly influenced by these factors so arbitrary from a moral point of view” (p. 72). Even if we adjust by introducing institutions that establish and sustain fair equality of opportunity, we still cannot endorse the distributive outcome of the competitive market economy as fair or at least not unfair. Of the system of natural liberty constrained by fair equality of opportunity, Rawls writes that “distributive shares are decided by the outcome of the natural lottery; and this outcome is arbitrary from a moral perspective” (p. 74).

What is going on here? For three different, interesting takes on this move in Rawls, see Robert Nozick, *Anarchy, State and Utopia*, chapter 7, the end of the chapter; Susan Hurley, *Justice, Luck, and Knowledge*; and G. A. Cohen, *Rescuing Justice and Equality*.

Notice first that under Rawls’s system of principles, in a society in which Rawlsian principles are fulfilled, presumably talented people will tend to win positions of advantage and enjoy unequal, better than average levels of income and other social goods, so the natural lottery plays a role in determining distributive shares, but one supposes this is not arbitrary from a moral perspective. Luck comes in different flavors, and which ones one finds morally objectionable is a matter that depends on the fundamental moral principles one accepts.

What is morally arbitrary about the natural lottery? It distributes natural assets, and for any individual, her holding of natural assets is a matter of brute luck—good or bad luck that falls on her in a way that is entirely beyond her power to control. So maybe the line to draw is between inequalities that arise by brute luck and inequalities that arise in other ways that are mediated by choice. The extent to which an outcome is brute luck or not will then vary by degree. Some have seen in Rawls’s remarks a presumption in favor of equality of condition—it is morally bad (unjust and unfair) if some are worse off than others through no fault or responsible choice of their own (the formulation is from Larry Temkin, *Inequality*).

But on that line, how do we get from a presumption of equality of condition to the difference principle? Rawls stipulates that a reasonable person should not be envious: one should care intrinsically only about the basket of goods one gets for oneself from social arrangements and not about how one’s basket compares with the baskets of goods others are getting. Rawls also supposes that from an initial equal distribution, one should prefer to move from equality to another distribution in which everyone is better off, even if some are better off than others. A further move: From an initial equality, one should not object to a proposal to move to a new arrangement that makes someone else better off without making anyone worse off even if the someone who is made better off is not you.

I mean only to flag an issue for thought and discussion here. Rawls’s difference principle is wrapped in the language of equality—the principle is formulated in answer to the question, under what conditions are inequalities morally permissible? But the difference principle actually states that the best just distribution is the one that makes the worst off group as well off as possible.

Equality and departures from equality are neither here nor there. The difference principle cares about maximizing from the standpoint of the worst off. The relations between one person's holdings and another's do not matter at all, from the standpoint of the difference principle. Equality is a relational principle and the difference principle is a non-relational principle.

Another question: Suppose we accept the intuitive idea that justice bids us to favor worse off people, or perhaps better, those whose level of benefits and burdens assessed on an absolute scale, is poor. This consideration identifies a large family of views—the prioritarian family. The difference principle is extreme prioritarian—maximize from the standpoint of the worst off. Why accept the extreme priority accorded the worst off that is built into the difference principle?? Consider this choice: one could either gain a penny for one worst off person or lose that penny but gain benefits of any magnitude (a move from mediocrity to permanent bliss) for any number, however large, of slightly better off people. (In chapter 3 Rawls says his maximin principle is not intended to be applied to such choices.)

Another question: As formulated, the difference principle does not say, maximize the opportunity for income and wealth and other social benefits of the worst off. It says, maximize the income and wealth and other social benefits going to the worst off. But what level of income I end up with depends not just on my initial prospects but also on my own choices and conduct throughout life. Some inequalities that come about in this way, say by my willful absenteeism from work or my taking only part-time employment, look to be, in the terminology of the “basic structure as subject” section, shallow not deep inequalities. Why is this so? Rawls has answers, we shall see. One is that the difference principle applies to groups not individuals.

6, Primary social goods as the basis of expectations. One further feature of Rawls's principles is their resourcist character. That is, the principles of justice require basic structural institutions to be set so they distribute resources to individuals in a way that the principles endorse. These resources Rawls class “primary social goods.” The idea undergoes some shifting as Rawls's thought develops over time. In the 1971 edition of *A Theory of Justice*, primary social goods are those goods, distributable by society, that any rational individual wants more of rather than less. Rawls supposes these goods are “rights and liberties, opportunities and powers, income and wealth.” (p. 93). There is also “a sense of one's own worth” or the social prerequisites for that. In later writings, Rawls alters the idea of a *primary social good*. Primary social goods are stipulated to be those social goods that any individual who gives priority to her interests in developing and exercising her capacity for a sense of justice (cooperating with others on fair terms) and for a conception of the good would want more of rather than less. Although the definition of “primary social good” shifts, Rawls seems to suppose the two different definitions pick out roughly the same goods.

The Equal Liberty Principle says that a class of primary social goods, the basic liberties should be arranged so that each person has a maximal share consistent with this same maximal share going to every other member of society. We would face measurement problems if we faced choices in which we could improve people's enjoyment of some basic liberties at a cost of lesser enjoyment of others.

Basic liberty does not trade off against other social and economic benefits, so the question how to weight basic liberties against these other benefits does not arise. But to identify the worst off group of persons in society we must be able to say, given a distribution across persons of the various primary social goods (other than basic liberties), who has how much of primary social goods overall and whose holding is the smallest. We need an index of primary social goods other than basic liberties, in other words. Some think this is a big problem for Rawls; others don't think so.

The fact that social justice assessments are to be made in terms of people's holdings or expectations of primary social goods marks a significant disagreement between Rawls and utilitarianism. For the utilitarian, the distribution of resources is a tool, to be set so that good

quality of life for individuals (happiness, desire satisfaction, achievement of excellence—depending on one's notion of utility) is maximized in the aggregate. One might think the utilitarian is right to focus on the good quality of life that the resource distribution generates rather than the resource distribution per se as of fundamental moral importance, even if one disagrees with the utilitarian's discounting of the importance of the fair distribution of good quality of life across persons.

Here is Rawls on this issue: "It may be objected that expectations should not be defined as an index of primary goods anyway but rather as the satisfaction to be expected when plans are executed using these goods. . . . Justice as fairness, however, takes a different view. For it does not look behind the use which persons make of the rights and opportunities available to them in order to measure, much less to maximize, the satisfactions they achieve, Nor does it try to evaluate the relative merits of different conceptions of the good. Instead it is assumed that the members of society are rational persons able to adjust their conceptions of the good to their situation" (section #16, page 94).

Rawls use of primary social goods as the basis of interpersonal comparisons for a theory of justice incorporates a division of moral labor between society and the individual and also a division of moral responsibility. Society (all of us regarded collectively) is responsible (obligated) to arrange institutions and practices so that each individual gets a fair bundle of primary social goods, the shape of the fair bundle being fixed by fundamental justice principles. Once these fair shares of resources are sustained, how one's life goes is up to each individual. Given a fair distribution of resources and a fair framework for social interaction, each individual makes her plans and tries to fulfill them and ends up with a better or worse quality of life. Against a fair background the individual is responsible for the life outcome she reaches, in the sense that she has no basis for demanding compensation or further resources from society on grounds of justice if her life does not go as well as it might have done. In this way the Rawlsian liberal just society is not a nanny state. (In this connection see Rawls's essay "Social Unity and Primary Goods" for his furthest development of this theme.