

**PHIL 167: Contemporary Political Philosophy**  
**Fall 2005; David O. Brink**  
**Handout #6: Liberal Egalitarianism II: Rawls on Justice as Fairness**

**JUSTICE AS FAIRNESS**

In A Theory of Justice (1971) John Rawls (1921-2002) develops an egalitarian conception of social justice which he defends by appeal to a hypothetical social contract. Traditional social contract theory answers the question

- Under what conditions is state authority justified?

by appeal to actual or hypothetical consent. Rawls's question is

- Under what conditions is the basic structure of society just?

His answer is that it is just iff it satisfies principles that would have been agreed to under fair conditions. Hence, Rawls's decision to call his theory Justice as Fairness. He must specify (a) fair initial contractual circumstances and (b) the principles that would be chosen in these circumstances.

**THE TWO PRINCIPLES OF JUSTICE**

Let's begin with principles the principles that Rawls thinks would be chosen in this hypothetical contract. There are two.

1. 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 [**Equal Basic Liberties**].
2. Social and economic inequalities are to be arranged so that they are both (a) to the greatest benefit of the least advantaged ... [the **Difference Principle**] and (b) attached to offices and positions open to all under conditions of fair equality of opportunity [**Fair Equality of Opportunity**].

The three principles here regulate the distribution of different primary goods. The Difference Principle is Rawls's only departure from equality. He thinks that it would be irrational for anyone to prefer social and economic equality to a system that allowed inequalities that maximize the position of the worst-off.

	A	B	C
D1	10	10	10
D2	12	13	14
D3	5	10	25

Suppose D1 is the maximum equal distribution. The Difference Principle would condone (indeed, require) D2. Utilitarianism would require D3, which both Equality and the Difference Principle would condemn.

Rawls takes utilitarianism and, to a lesser extent, libertarianism as the chief rivals to his more egalitarian conception of justice. Some of Rawls's arguments are arguments of overall comparative plausibility, independent of the contract argument. The libertarian conception gives moral significance, as justice as fairness does not, to the morally arbitrary effects of the social and natural lotteries (TJ 65-83). Laissez-faire must give way to a scheme which recognizes the moral arbitrariness of the social lottery and, ultimately, to one which recognizes the moral arbitrariness of the natural lottery as well. In contrast with utilitarianism, Justice as Fairness places a lower limit on the sacrifices that some might have to make in order to improve the position of others (28, 175-83, 500).

### **THE CONTRACTUAL ARGUMENT FROM THE ORIGINAL POSITION**

But Rawls's most distinctive argument for his conception of social justice is his social contract argument. Rawls's social contract is hypothetical (12, 121), appealing to an **Original Position** that excludes factors arbitrary from the moral point of view (120). Parties in the original position (OPs) are placed behind a veil of ignorance, which deprives them of information about their identities and attributes (137), including

- identity
- gender
- race
- class position
- natural assets
- aversion to risk
- conception of the good
- circumstances of society
- generation

However, OPs do have positive characteristics. They

- are mutually disinterested, but concerned to advance their own prospects (143)
- use primary goods (maximally flexible social assets) to measure their prospects (62, 92)
- evaluate representative social positions (64)
- display instrumental rationality (143)
- have general social knowledge (137)

This presents the following idealized choice situation: Knowing all of the possible representative social positions that she might occupy in all possible societies at all possible points in time, an OP should choose principles of justice that set terms of social cooperation that will best promote her stock of primary goods.

Though OPs are not making a moral choice, they realize the point of their exercise and assess the consequences of alternative principles by considering their application in a well ordered society (TJ: 4-5, 453-62). In a **well ordered society**, citizens have a sense of justice and a higher order interest in pursuing their conception of the good, whatever it is, once the veil of ignorance is lifted (TJ: 4-5, 145). Moreover, principles must satisfy a **publicity** condition, according to which the principles are recognized publicly as establishing standards of just institutional design and citizens actually assess their institutions in private and public fora by

measuring their conformity to these principles (TJ: 6, 133). Publicity, Rawls believes, imposes constraints of **stability** and **practicality** (TJ: 138, 140, 142).

### **MORE ABOUT THE TWO PRINCIPLES**

The two principles, like their rivals, apply to the basic structure of society. The two principles constitute what Rawls calls the **Special Conception** of justice (TJ: 302-3):

1. Each person is to have an equal right to the most extensive total system of basic liberties compatible with a similar system of liberty for all.
2. Social and economic inequalities are to be arranged so that they are both (a) to the greatest benefit of the least advantaged, consistent with the just savings principle, and (b) attached to offices and positions open to all under conditions of fair equality of opportunity.

(1) and (2) regulate the distribution of different primary goods.

The Equal Liberties principle concerns specific basic liberties, rather than liberty per se. As Hart points out, the most extensive system of equal liberty would not permit restrictions on liberty to prevent harm or nuisance. Because this is not Rawls's claim, he must be concerned with specific liberties that seem fundamental. The first principle concerns those liberties relevant to the two primary capacities of citizens in a well ordered society: (a) their sense of justice and (b) their interest in pursuing a conception of the good.

Rawls's initial characterization of the second principle requires only that opportunities be equal and that inequalities be to everyone's advantage. Rawls goes on to consider three different interpretations of this initial claim (§§12-13). The system of natural liberty requires that "careers open to talents" and circumstances and accepts distributions resulting from efficient market exchanges. But the liberal conception objects that the system of natural liberty allows inequalities that are produced by effects of the **social lottery**. It insists on an interpretation of fair equality of opportunity that corrects for effects of the social lottery. But the **natural lottery** is equally arbitrary from the moral point of view. The conception of democratic equality corrects for effects of the natural lottery as well by treating natural talents as **common asset**. This, Rawls claims leads one to his second principle, in particular, the Difference Principle.

As we saw, the Difference Principle treats the (highest) equal distribution as default for purposes of assessing potential inequalities.

	A	B	C
D1	8	8	8
D2	12	13	14
D3	9	15	18

It prefers D2 to D1. Notice that the Difference Principle is not the same as Trickle Down, which only requires that inequalities make all boats rise. D3 satisfies Trickle Down, but the Difference Principle requires D2, rather than D3.

Where two or more distributions maximize the worst-off, presumably the Difference Principle directs us to then maximize the position of the next worst-off, and so on.

	A	B	C
D1	5	7	11
D2	5	8	10
D3	5	9	9

If so, the Difference Principle would prefer D2 to D1 and D3 to D2. This is to treat the Difference Principle as equivalent to Leximin, rather than Maximin.

Notice also that the "worst-off" has **variable reference**, possibly referring to different individuals or groups of individuals under different distributions (i.e. functions as a definite description, rather than a rigid designator).

	A	B	C
D1	5	7	9
D2	8	8	7
D3	9	7	5

So, if D1 is the actual distribution, then A is the worst-off in the actual world. The Difference Principle requires preferring D2 to D1, even though now C will be the worst-off. The Difference Principle does not require D3, which would maximize A's position (in fact, it is indifferent between D1 and D3). Though Rawls generally recognizes this fact about the difference principle, he seems to forget it at one crucial point in the justification of the Difference Principle (TJ 103).

The Difference Principle does treat natural talents as a common asset (TJ: 101). But notice that this does not imply, as Nozick believes (ASU 225, 229), that the community has property rights in the talents and powers of individuals, such that the community could compel talented individuals to work on the community's terms.

1. I am entitled to possess my natural endowments, though they are unearned.
2. I am entitled to exercise my natural endowments (in acceptable ways).
3. I am entitled to benefit from the exercise of my natural endowments (assuming they're productively employed).
4. I am entitled to all the benefits (that I can get others to concede to me) from my exercise of my natural endowments.

Rawls does not deny (1)-(3), only (4).

Notice Rawls's priority rules. (1) is "lexically prior" to (2), and (2b) is "lexically prior" to (2a) (TJ 302). Also, Rawls distinguishes between having certain rights or liberties and the **worth** or **value** of those rights or liberties. You and I may have the same rights to contribute money to the campaign of our favorite political candidates, but your right or liberty is worth more to you than mine is to me if you have significantly more resources than I do. Rawls insists not only on equal basic liberties but also on the **fair value of political liberties**. The fair value of political liberties is "lexically prior" to (2), especially (2a) (TJ 204, 225-26, 315-16).

We have been examining Rawls's Special Conception, which applies in societies above a certain developmental threshold, at which point Rawls thinks individuals care more for marginal increases in basic liberties than for marginal increases in social and economic goods. Below this

developmental threshold, Rawls endorses the **General Conception**, which basically distributes all goods according to the Difference Principle.

All social primary goods -- liberty and opportunity, income and wealth, and the bases of self-respect -- are to be distributed equally unless an unequal distribution of any or all of these goods is to the advantage of the least favored [TJ 303].

Notice that the Special Conception differs from the General Conception in two ways. (a) It treats the distribution of some goods as lexically prior to others, and (b) it insists on an equal distribution for liberties and opportunities, while requiring that other goods be distributed according to the Difference Principle. We should assess these two points of difference separately. Rawls's claim about people caring more for liberties than other goods above a certain threshold of material well-being seems only to address (a), not (b). But then, even if we accept (a), we may not accept (b). Just as Rawls thinks that it's irrational to prefer Equality to the Difference Principle when considering other primary goods, why not prefer the Difference Principle to Equality for liberties and opportunities?

Perhaps we should give liberties lexical priority over opportunities and other goods but distribute each by the Difference Principle. As it stands, Rawls defends

1. Equality for liberties.
2. Equality for opportunities.
3. Difference Principle for other primary goods.

But maybe he should endorse

1. Difference Principle for liberties.
2. Difference Principle for opportunities.
3. Difference Principle for other primary goods.

Moreover, we might question (a). If there is a preference, above a certain threshold, for liberties over opportunities or other primary goods, surely it is not a lexical preference. Even if I wouldn't trade increases in basic liberties one-for-one with increases in other primary goods, surely I might prefer large increases in other primary goods to small increases in basic liberties. I might prefer to live much more comfortably, even if this is purchased at a very small cost in my personal or political liberty. But if we reject Rawls' egalitarianism about liberties and opportunities and reject his lexical priority rules, the result is something closer to the General Conception, even for societies above the material threshold.

## **APPLYING THE TWO PRINCIPLES**

Justice as Fairness is to be applied in a four stage sequence, in which the veil of ignorance is gradually lifted (TJ§ 31).

1. OP choice made behind a veil of ignorance.
2. Constitutional stage in which OPs know which society they live in in which historical conditions with which natural and social resources.
3. Legislative stage at which legislators design institutions and policies and seem ignorant only about who occupies what positions.
4. Judicial stage at which constitutional and statutory provisions are applied at which point the veil of ignorance seems to be completely lifted.

Choices made at any stage are constrained by choices made at prior stages.

### THE ARGUMENT AGAINST UTILITARIANISM

Within the contractual argument, Rawls takes utilitarianism to be the main rival to justice as fairness. We must distinguish between **classical utilitarianism**, which tells agents to maximize total happiness, and **average utilitarianism**, which tells agents to maximize average or per capita happiness.

	A	B	C	D
D1	5	5	5	
D2	4	4	4	4

Whereas classical utilitarianism endorses D2, average utilitarianism endorses D1. OPs would prefer average to classical utilitarianism, because it maximizes their expected prospects.

### THE CONTRACT-INDEPENDENT ARGUMENT

- Utilitarianism seeks to maximize the good and, hence, attaches no intrinsic significance to the distribution of goods (TJ 26).
  - The utilitarian can often appeal to **diminishing marginal utility** to condemn inequalities or promote equality. But this attaches instrumental, rather than intrinsic, significance to distributional issues. This is illustrated by the possibility of "pleasure wizards" and "utility black holes" for whom diminishing returns set in very late, if at all.
- Utilitarianism ignores the **separateness of persons** (TJ 26-31). Utilitarianism tell us to balance benefits and harms both within and across lives, where necessary, in order to secure the best outcome overall. The sort of interpersonal balancing that utilitarianism requires is like the sort of intrapersonal balancing that prudence requires. Prudence requires a person to be concerned about her own overall good and to balance benefits and harms at different points in her life so as to produce the most happiness overall. Just as prudence is impartial or neutral about the temporal location of benefits and harms within an agent's life, so too utilitarianism is impartial or neutral about the distribution of benefits and harms across lives. But whereas it may seem acceptable to balance goods and harms **within** a life, as prudence requires, balancing goods and harms **across** lives appears to ignore the separateness of persons. Sacrifices made in the course of intrapersonal balancing are automatically **compensated**, because benefactor and beneficiary are one and the same person. However, sacrifices made in the course of interpersonal balancing are not automatically compensated, because benefactor and beneficiary are different people. Rawls, like Nozick, insists that the separateness of

persons gives individuals **rights** that should trump or defeat considerations of collective good (TJ 27-28;cf. ASU 28-30). But we might raise at least two issues in reply.

- The separateness of persons objection rests on the contrast between intrapersonal and interpersonal unity. We could question this contrast if we think that different people can be and are related to each other in ways like the ways in which different stages or periods within a life can be and are related to each other. For instance, Derek Parfit defends a form of **psychological reductionism** about personal identity that seems to imply that the separateness of persons involves differences of degree, not kind (Reasons and Persons, part III). This view allows us both to see intrapersonal disunity and interpersonal unity. However, we need to be clear how we assimilate the intrapersonal and interpersonal cases.
  - (a) We might assimilate different periods within a life to the way we had viewed different lives by noting how earlier and later versions of myself can appear foreign or other to each other. Assimilating the intrapersonal to the interpersonal would argue for **extending the scope** of distributional principles from interpersonal contexts into intrapersonal contexts. On this view we need intrapersonal, as well as interpersonal, distributive justice.
  - (b) Alternatively, we might assimilate interpersonal relations to intrapersonal relations by noting that the sort of causal influence between different periods within a single life can be found between lives as well. But if we assimilate the interpersonal to the intrapersonal, we might **reduce the weight** of distributional considerations.

(b), not (a), provides some support for utilitarianism (cf. Parfit, RP, ch. 15). However, even (b) does not deliver utilitarianism. I am more closely connected psychologically with some people than with others (e.g. with intimates than with strangers). But if interpersonal regard should track these psychological relations, this would support an interpersonal discount rate, rather than utilitarianism's person neutrality.

- Because Rawls criticizes utilitarianism for demanding uncompensated sacrifices, we might assume that the Difference Principle does not require uncompensated sacrifices. But this is not exactly right. Rawls asks us to compare the sacrifices of the better-off (A) and the worse-off (B).

Now B can accept A's being better off since A's advantages have been gained in ways that improve B's prospects. If A were not allowed his better position, B would be even worse off than he is. The difficulty [for the Difference Principle] is to show that A has no grounds for complaint. Perhaps he is required to have less than he might since his having more would result in some loss to B. Now what can be said to the more favored man? To begin with, it is clear that the well-being of each depends on a scheme of social cooperation without which no one could have a satisfactory life. Secondly, we can ask for the willing cooperation of everyone only if the terms of the scheme are reasonable. The difference principle, then, seems to be a fair basis on which those better endowed, or more fortunate in their social circumstances, could expect others to collaborate with them when some workable arrangement is a necessary condition of the good of all [TJ 103].

(Notice that Rawls here treats "better-off" and "worse-off" as rigid designators, rather than definite descriptions, contrary to his official position.) As Nozick notes (ASU 192-96), the better-off can complain that they are required to accept lower prospects so that the worse-off can enjoy higher prospects. Rawls has not yet explained any asymmetry between the sacrifices of the better-off and worse-off. The cooperative surplus depends on the willing cooperation of both. Rawls must be assuming that the better-off have no moral claim to a large share of the cooperative surplus that their natural talents help make possible, that they are entitled to benefit from the productive employment of the natural talents only on morally acceptable terms, and that **all else being equal bottom-up sacrifice is morally worse than top-down sacrifice**.

- Utilitarianism violates stability and the strains of commitment. A conception of justice is stable when its public recognition and implementation tends to support, rather than undermine, continued adherence to that conception. Because utilitarianism might require some to sacrifice in order to promote total (or average) utility, utilitarianism could only be a stable conception if people were willing to identify themselves with others to an unnatural (and unwholesome) extent (TJ 178). By contrast, Rawls believes that the Difference Principle provides a stable basis for social cooperation (TJ 103). However, as we have seen, one might wonder whether the Difference Principle is similarly unstable because it requires some (the better-off) to accept lower prospects for the sake of others (the worse-off) (cf. ASU 192-96). Perhaps Rawls believes that there is a moral asymmetry between bottom-up and top-down sacrifice.

### THE CONTRACTUAL ARGUMENT

Rawls claims that OPs would prefer his two principles to (average) utilitarianism. He contrasts the maximin decision rule with the rule of maximizing expected utility. These two decision rules will evaluate possible distributions differently.

	A	B	C
D1	10	15	20
D2	12	13	15
D3	5	15	26

Assume that this matrix represents possible distributions of utility (rather than primary goods). Should OPs maximize expected utility and gamble that they will be one of the better-off in D1 or D3 or should they maximize the worst case scenario and play it safe with D2? Rawls says there are three conditions that favor maximin (TJ 154-55). (Actually, it would be better if OPs would employ leximin, rather than maximin.)

1. Ignorance of the probabilities of the various outcomes.
2. A conception of the good such that one cares little, if anything, for what one might gain above a certain minimum.
3. Alternative decision rules have possible outcomes that would be unacceptable.

Rawls claims that these three conditions are satisfied or approximated in the original position (TJ 155-56). In assessing this argument, we need to distinguish between (a) whether these

conditions, individually or collectively, favor maximin, and (b) whether these conditions are satisfied or even approximated in the OP.

(1a) Decision theory distinguishes decision under **certainty**, **risk** (which includes probability spaces, as well as determinate probabilities), and **uncertainty** or **ignorance**. It identifies rational decision making with maximizing expected utility (MEU). MEU is straightforward in contexts of certainty (prefer more utility to less) and in contexts of risk (prefer 1/3 chance of 100 utiles to a 1/2 chance of 50 utiles to a 1/1 chance of 20 utiles). It is usually thought that rational decision making is impossible and/or undefined for uncertainty or ignorance. So I think that the usual view would be that maximin is not a rational decision principle under any circumstances; it loses out to MEU in contexts of certainty and risk, and neither principle is rational for ignorance.

(1b) Rawls claims that the OP does involve ignorance of probabilities. Is this appropriate? Should OPs be ignorant of the probabilities of whether they will be A, B, or C, or should they assume that these possibilities are equi-probable?

(2a) As it stands, (2) does not help the case for maximin. The conception of the good must exhibit a marginal utility threshold above which one cares very little or nothing for further increments in primary goods that is at or below the minimum that maximin would secure. Moreover, above this threshold, one must care absolutely nothing for further increments of primary goods. So (2) must really read as follows

(2') A conception of the good such that one cares nothing at all for increments above the highest minimum that one might receive.

Two points about (2'). First, it's no surprise that it supports maximin. Second, it is a very extreme assumption.

(2b) OPs do not know their conceptions of the good. A fortiori they do not know that they have this peculiar conception of the good.

(3a) Of course, if all the outcomes under alternative decision rules were unacceptable or intolerable, then this would be a clear argument in favor of maximin. But that is not true. So what follows from the fact that some outcomes under MEU would be worse than under maximin but some would be better?

(3b) I think that a large part of Rawls's case for maximin depends on the idea that the stakes are so high in the OP. This is not a small gamble; it is a single choice that determines one's entire life prospects. The costs of losing one's shirt in this gamble are enormous. In this context, Rawls thinks that it really would be literally intolerable to accept some possible results under MEU (e.g. slaving away in a utility maximizing death camp so that the commandants who are utility black holes can have a high time). Maybe the worst-case in D1 wouldn't be intolerable, but D3 is a possible outcome, and its worst-case might seem intolerable, especially when in D3 A remembers that D2 was a safety net which he refused. This is perhaps Rawls's strongest argument. However, we want to make sure that he is not in fact reassigning utilities in considering D3, creating a D4 in which A has infinitely low or negative utility. For MEU might condemn D4 as well.

In any case, these arguments do not take Rawls far enough in defending the Difference Principle.

- Rawls wants to defend the Difference Principle not just against unrestricted (average) utilitarianism, but also against so-called **mixed conceptions** that, like Justice as Fairness,

distribute social and economic goods only after ensuring equal basic liberties, the fair value of political liberties, and equal opportunity. But he argues only against unrestricted utilitarianism. Rawls illustrates the claim that utilitarianism is compatible with intolerable outcomes by appealing to a social system in which the better-off secure great social and economic advantages by restricting the basic liberties of the worse-off, as in systems of slavery (TJ 156). But utilitarianism would not allow this if it figured in a mixed conception, for it would allow interpersonal balancing only after equal basic liberties, the fair value of political liberties, and equal opportunity had been secured. These **prior principles** will significantly constrain the range of admissible inequalities. But if utilitarianism figures in such a mixed conception, it seems very unlikely that it can have intolerable outcomes. Perhaps utilitarian gambling above this sort of safety net is rational and/or attractive.

- Or, if ensuring equal basic liberties, the fair value of political liberties, and equal opportunity does not already rule out intolerable outcome, posit a tolerability threshold above this, and consider a mixed conception that employs MEU for all and only distributions above the threshold. Call this view **Sufficientarianism**. As long as that threshold is below the maximal minimum, Sufficientarianism will be different from the Difference Principle. It's hard to see what Rawls's argument could be that OPs would rationally prefer the Difference Principle to Sufficientarianism.