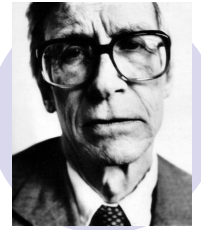# RAWLS

In a hypothetical choice situation modeling fairness, we'd agree to principles of justice ensuring basic liberties and allowing inequalities only where they benefit the worst off.

Aims of Rawls's theory

- **to formulate and defend principles of distributive justice governing the basic structure of [a democratic pluralist] society – vs. particular social programs or rules of interpersonal morality**

- **to disallow utilitarian trade-offs (the sacrifice of some people's interests for the greater overall good), by recognizing the separateness of individual persons**

- **to provide the basis for a "well-ordered society," known to be governed by principles of justice by members who exhibit a corresponding sense of justice.**

  - **not a utopia: correction, rather than elimination, of unjust acts: sense of justice tends to bring people back into line when they violate the principles**

  - **an ideal society: not what we're like now, but what we'd be like if we were brought up to share the conception of justice corresponding to the principles of justice – not to be confused with the hypothetical situation resulting in the choice of those principles (see next slide)**
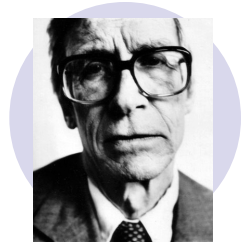
The "original position"
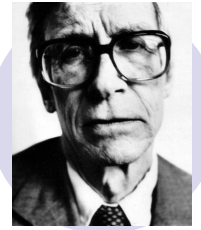
= a hypothetical situation embodying "justice as fairness," the basis for deriving principles of justice

- ○ not a conception of human nature or society, actual *or* ideal, or of what we'd be like if we hadn't yet set up a system of justice — though it plays a role parallel to the "state of nature" in historical contract theories

- ○ defined by two main conditions:

    - ● <u>"the veil of ignorance"</u>: Parties to the contract lack knowledge of facts about themselves or their societies in particular, e.g. their place in society, native endowments, psychology, conception of the good.

    - ● <u>rational self-interest</u>: Each chooses what's best for himself, without any independent concern for others' interests.

Rawls's basic argument

- **Rawls argues from the original position (OP) to two principles of justice roughly as follows:**

  1. **The OP represents a fair choice situation among free and equal persons.**

  2. **It would be rational for individuals in the OP to agree to set up the basic structure of society on the basis of certain principles.**

  3. **So those principles must be fair and just.**

- **To confirm the first premise, Rawls relies on his method of "reflective equilibrium": mutual adjustment of our assumptions about the OP and our intuitions as to which principles are fair – in comparison with traditional alternative principles, especially utilitarianism, which applies individual rationality to society as a whole.**

- **Rawls's version of the second premise rests on an analogue of the principle of "maximin" from rational choice theory, which tells you to "maximize the minimum," i.e. to make your worst outcome as good as possible.**

Progression of Rawlsian concepts
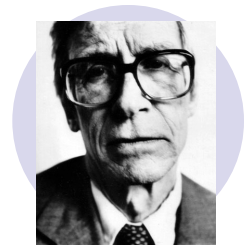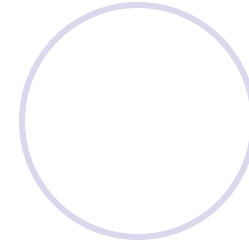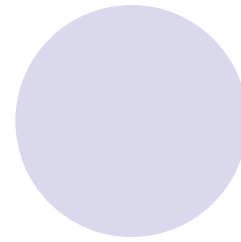
intuitions on fairness

[mutually adjusted by method of reflective equilibrium]

Original Position

[rational choice according to maximin]

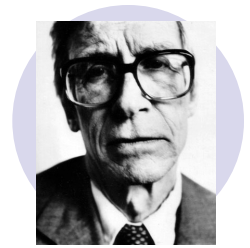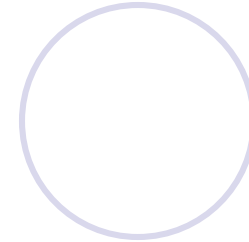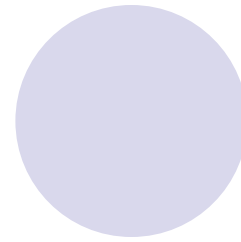## Principles of Justice

[put into force and made public]

Well-Ordered Society

[development of moral sentiments into sense of justice, supporting stability]

The two principles of justice (1)

- Rawls argues that parties to a contract in his OP would agree unanimously to arrange the basic structure of society according to two principles; see p. 266 for their final form (cf. p. 54 for the general idea behind them).

- Together the principles should ensure that the worst outcome for any individual is as good as possible without restricting anyone's freedom to pursue a particular conception of the good.

- The two principles are in serial ("lexical") order, i.e. the first must always be satisfied before the second.

- The first principle can thus be summed up roughly as:

1. the priority of liberty: Everyone has an equal right to the maximum basic liberties [see list, p. 53] compatible with like liberties for all.

The two principles of justice (2)
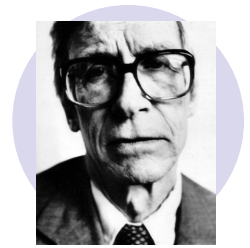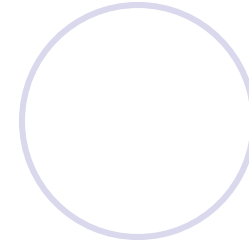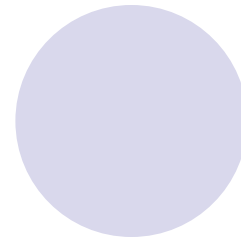
- The second principle is more complex. In simplified form:

2. inequalities in distribution of primary goods (see list, p. 54) must satisfy two conditions:

  a) <u>the difference principle:</u>   benefit the worst off [and thereby benefit everyone]

  b) <u>fair equality of opportunity</u>: attached to positions and offices open to all

- Here the second part always overrides the first. Rawls maintains that the ordering of the two principles and their parts prevents *extreme* inequalities from being justified by the difference principle, since the priority of liberty plus fair equality of opportunity has results for education and the like with a "leveling" effect. The difference principle comes in last, to demand justification for any remaining inequalities.

Rationality assumptions
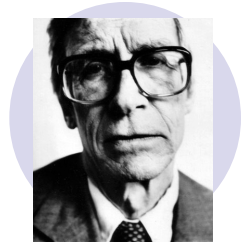
- **Rationality sometimes requires accepting inequalities, on the assumption that envy is irrational. Only your *absolute* welfare should matter, but envy involves wishing others were worse off even if that would make you worse off too.**

- **Maximin is rational under three conditions that Rawls takes to apply to the OP:**

  1. **The parties can't assign definite probabilities to the different possible outcomes (because they don't know facts about their particular societies, e.g. how large each socioeconomic class is).**

  2. **They wouldn't lose that much if they don't achieve greater than the minimum benefit (because inequalities are minimized by the priority of liberty and fair equality of opportunity).**

  3. **They'd lose a huge amount if they didn't achieve a certain minimum.**

Sample distributions
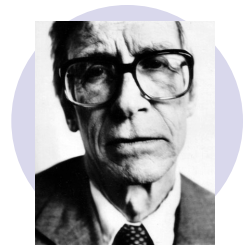
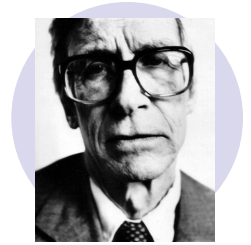|	Best off	Worst off	Total
Strict Egalitarian	80	80	160
Rawlsian	**150**	**100**	**250**
Utilitarian (1)	200	70	270
Utilitarian (2)	170	100	270

Objections to Rawls's principles

- **Rawls has to deal with objections from both ends of the political spectrum, as well as from utilitarians and other philosophers.**

  ○ **Many on the left would object to the priority of liberty and departures from strict egalitarianism.**

  ○ **Many on the right would object to compromising even non-basic liberties on the basis of the difference principle.**

- **Here's a sketch of some non-utilitarian objections within philosophy (along with an indication of possible replies):**

  ○ **Doesn't it matter how a given distribution came about? Cf. Nozick's Lockean model of "justice in acquisition/transfer." (But how can we correct historical contingencies that may have been arbitrary or unjust?)**

  ○ **Isn't Rawls really capturing justice *plus* something else, e.g. benevolence, that doesn't count as "the first virtue of social institutions," or something to which one has a right? Cf. Nagel on Rawls's insistence on correcting the "natural lottery." (But can we expect people to adhere to a social system that leaves them with disadvantages they weren't responsible for?)**

Rawls's main alternatives
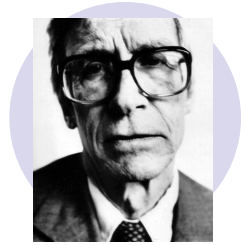
- **<u>Utilitarianism:</u>**

  - ○ **Rawls's main contrast is to utilitarianism, which he conceives as based on a single rational individual identifying sympathetically with all sentient beings lumped together.**

  - ○ **As such, it asks what would promote the advantage of everyone as a whole, ignoring the separateness of individuals and thus allowing trade-offs.**

- **<u>Perfectionism:</u>**

  - ○ **The other option Rawls considers is perfectionism, which advocates maximizing individual excellence – a view he attributes to Aristotle and Nietzsche.**

  - ○ **When he later discusses self-respect as a primary good, he views it as properly extending beyond persons who achieve all-round excellence, as opposed to excellence in using their particular talents and abilities to follow their own rational-life plans in groups of supportive individuals..**

The Well-Ordered Society

- **The ideal emerging from Rawls's theory is a society designed to advance the good of its members, ordered in accordance with the two principles of justice, and satisfying two conditions:**

  - **publicity: Everyone knows and accepts the principles and knows that others do too.**

  - **stability: Citizens acquire the corresponding sense of justice and are motivated to do their part in maintaining its institutions.**

- **Stability depends partly on generating internal forces to correct departures from justice, which requires that people have the appropriate moral sentiments.**

- **So while Rawls follows Kant in appealing only to rationality as the basis for his system, he needs an account of the development of moral sentiments capable of maintaining it.**

Moral psychology
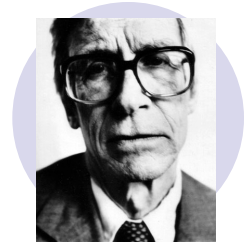
- **Rawls's account is meant to combine elements of two traditions of thought in philosophy and psychology that he sketches as follows:**

    - **<u>empiricist</u> (e.g., Hume, social learning theory, Freud): Moral motivation has to be learned, via reward and punishment, at a childhood stage preceding understanding.**

    - **<u>rationalist</u> (e.g., Rousseau, Kant, Piaget, sometimes Mill): Our natural sympathy provides an innate basis for moral motivation that develops further with the emergence of adult understanding.**

- **His Hume/Mill contrast may seem odd, since both are empiricists with utilitarian leanings, and Hume certainly stresses natural sympathy. But cf. Hume on our natural avidity, Mill on our need to be in harmony with others.**

The Morality of Authority
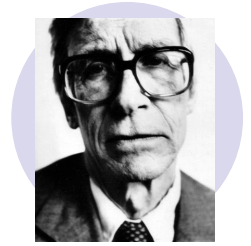
- **The first stage of moral development, before the child develops much understanding, involves:**

  - uncritical imitation of caretakers (including their responses to the child's behavior) as the source of moral attitudes

  - boosting of the child's self-esteem, which engenders love and trust for its caretakers, plus a form of guilt for infractions, as a result of the caretakers' evident love and care and adherence to their own precepts

- **See pp. 406 and 429 for what Rawls represents as his first "law of reciprocity."**

- **[Does this rule out development of moral sentiments on the part of children of unloving or unreliable caretakers?]**

The Morality of Association
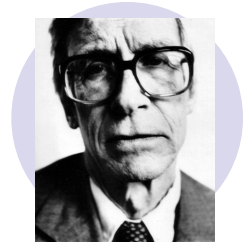
- **The second stage extends from family relations to membership in society and involves:**

  - mutual performance of group roles as source of moral attitudes

  - enough cognitive development for perspective-taking, which engenders "fellow feeling," plus a corresponding kind of guilt at nonperformance of one's own role within the group

- **See pp. 411 and 429 for Rawls's second law of reciprocity.**

- **There's also a role at this stage for emulation of exemplars of role-performance in the group – the sort of thing stressed by virtue ethics.**

The Morality of Principles
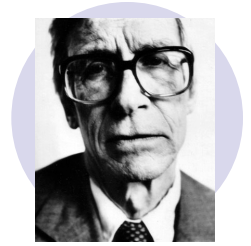
- **The third and final stage involves:**

  - **recognition of benefits from principles of justice as source of the sense of justice [= a kind of gratitude toward the system?]**

  - **development of guilt in the strict sense, for injustice toward *anyone*, not just those for whom one has "fellow feeling."**

- **See pp. 414f. and p. 429 for Rawls's third law of reciprocity.**

- **Moral guilt can be compounded by association guilt, for a worse infraction [again, in answer to contemporary virtue ethicists or others stressing personal relationships].**

- **Motivation to adhere to the principles is explained by the attractiveness of its underlying Kantian ideal, of cooperation between free and equal persons [= a kind of friendship?].**

Summary of the stages
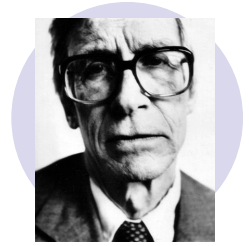
1. **morality of authority** (cf. first law of reciprocity, p. 406, p. 429):

   ○ uncritical imitation of caretakers as source of moral attitudes

   ○ caretakers' love and care engenders love and trust in the child, along with a form of guilt for infractions, by bolstering self-esteem.

2. **morality of association** (cf. second law, p. 411, p. 429):

   ○ mutual performance of group roles as source of moral attitudes

   ○ perspective-taking engenders "fellow feeling"/guilt at nonperformance

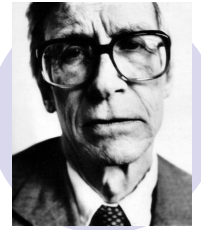3. **morality of principles** (cf. third law, pp. 414f., p. 429):

   ○ benefits from principles of justice as source of the sense of justice

   ○ development of guilt in the strict for injustice toward *anyone*
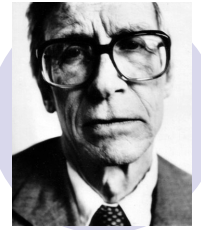
Comparison with utilitarianism

- **Rawls argues that a society based on his principles of justice would be more stable than one ordered in accordance with utilitarianism.**

- **In contrast to Rawls's reliance on *reciprocity*, utilitarianism would require *sympathetic identification* with society as a whole.**

  - **But those who can see that their welfare has been sacrificed for the sake of others are unlikely to develop attachments to and friendly feelings toward others who are better off.**

  - **Awareness of being used as a means to others' benefit (cf. Kant) would also tend to undermine their self-esteem, which is likely to diminish further their capacity for sympathetic identification with others.**

- **Rawls views his argument for the greater stability of his own system as answering the "free rider" egoist (cf. Hume's "sensible knave"). [But how does it apply to those who are better off? Would they also see themselves as benefitted by the principles and therefore abstain from violating them? Rawls would say yes, insofar as they imagine themselves in his OP.]**

Some rough historical contrasts

- **In his developmental account, as well as elsewhere, Rawls takes a lot from historical philosophers, including Hume and Mill.**

- **But it's worth noting some differences, e.g.:**

  - **For Hume, having a sense of justice as a moral virtue depends on extending our natural sympathy to society as a whole, once we recognize the general benefit of maintaining whatever system emerged from our mutual self-interest. Rawls instead bases the sense of justice on *reciprocity*, once we recognize that the particular system based on his principles benefits us and those we care about.**

  - **Though Rawls finds passages in Mill's ch. 3 to support his approach – Mill's emphasis on developing the sense that our individual good depends on that of others – in ch. 5 Mill claims that the sense of justice depends on "moralizing" retributive sentiments by applying the principle of utility. Rawls instead avoids basing the sense of justice on punitive motives and sees morality as developing out of reciprocity for benefits.**

Rawls's central argument

- **Rawls stresses that his treatment of the moral sentiments in defense of stability should not be taken as his argument for the two principles of justice. [Instead, it amounts to confirming evidence.]**

- **His argument rested on the application of his method of reflective equilibrium to the original position.**

- **As such, it depended on our intuitions – i.e. the considered intuitive judgments of competent judges, as defined in an earlier article, "Outline of a Decision Procedure for Ethics."**

- **Others have used Rawls's method to attack utilitarian judgments about cases other than justice – e.g., trolley cases, where the agent has a choice between killing one person and letting five be killed by a runaway trolley.**

- **However, there are contemporary objections to his method's reliance on intuitions.**

Questioning intuitions

- **A number of contemporary philosophers have argued that recent work in empirical fields undermines reliance on intuition as the basis for moral judgments and theories.**

- **As the term is used by most contemporary philosophers, following Rawls, moral intuitions amount to pre-theoretical judgments of right and wrong, etc.**

- **In an early piece on the subject that I've put online, Singer argues that we should mistrust our intuitions in constructing moral theories and instead rely on reasoning from the principle of utility.**

- **Besides rejecting Rawls's method, this means discounting apparent counterexamples to utilitarianism that rest on our intuitive objections to sacrificing one person's good to the good of a larger group.**

Sources of evidence

- **Singer's argument focuses on Rawls's method of reflective equilibrium and includes an overview of evidence from three main empirical sources. Briefly:**

  - <u>**evolutionary theory**</u> **(explanations of moral emotions by our evolutionary environment rather than current circumstances; see next slide)**

  - <u>**psychology**</u> **(Haidt's experiments purporting to show that ordinary moral judgment is formed intuitively, with reasons for it "confabulated" later)**

  - <u>**neuroscience**</u> **(Greene's brain-imaging studies linking deontological judgments to emotion, whereas consequentialist judgments emphasize slower processes of reasoning)**

- **[There's also evidence in the "heuristics and biases" literature of the influence of morally irrelevant factors such as order of presentation on moral judgments.]**

[Rawlsian intuitions?

- **A general problem for Singer's argument is that the empirical work he cites as evidence interprets "intuition" differently from Rawls.**

  - **What psychologists and others in empirical fields mean by "intuition" is something pre-reflective: responses not based on any conscious reasoning.**

  - **But Rawls specifies "considered" intuitions as input to reflective equilibrium. When he introduces his method, he says they can be based on everyday reasoning – e.g., factual inquiry, reflection on consequences, application of a common-sense rule – but just not conscious and systematic derivation from general theoretical principles.**

- **Even if our pre-reflective intuitions are evolved emotional responses, then, our pre-theoretical but "considered" intuitions might be based on correcting them, e.g. for changes from the evolutionary environment that don't seem morally relevant.]**