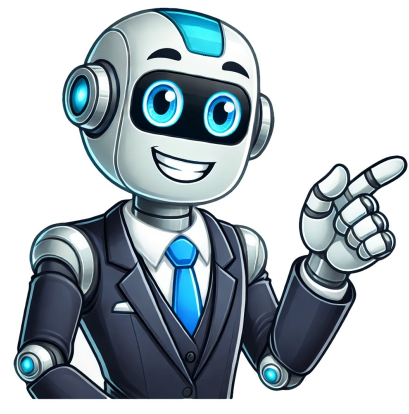


[Click Here](#)





































published Sat Dec 20, 2008; substantive revision Tue Oct 24, 2023 The idea of the moral point of view can be traced back to David Hume's account of the "judicious spectator." Hume sought to explain how moral judgments of approval and disapproval are possible given that people normally are focused on achieving their own interests and concerns. He conjectured that in making moral judgments individuals abstract in imagination from their own interests and adopt an impartial point of view from which they assess the effects of their own and others' actions on the interests of everyone. Since, according to Hume, we can all adopt this impartial perspective in imagination, it accounts for our agreement in moral judgments (See Hume 1739 [1978, 581]; Rawls, LHPM 84-93, LHPH 184-187). Subsequently, philosophers posited similar perspectives for moral reasoning designed to yield impartial judgments once individuals abstract from their own aims and interests and assess situations from an impartial point of view. But rather than being mainly explanatory of moral judgments like Hume's "judicious spectator," the role of these impartial perspectives is to serve as a basis from which to assess and justify moral rules and principles. Kant's categorical imperative procedure, Adam Smith's "impartial spectator," and Sidgwick's "point of view of the universe" are all different versions of the moral point of view. An important feature of the moral point of view is that it is designed to represent what is essential to the activity of moral reasoning. For example, Kant's categorical imperative is envisioned as a point of view any reasonable morally motivated person can adopt in deliberating about what they ought morally to do (Rawls, CP 339). Similarly, Smith's impartial spectator is a point of view that we can adopt in assessing the propriety of our actions for the good of the world (Smith, 1792 [1978, 10]). Hume's account of the moral point of view regarding matters of justice. The original position is a hypothetical perspective that we can adopt in our moral reasoning about the most basic principles of social and political justice. What primarily distinguishes Rawls's impartial perspective from its antecedents (in Hume, Smith, Kant, etc.) is that, rather than representing the judgment of one person, it is conceived socially, as a general agreement by (representatives of) all adult members of an ongoing society. The point of view of justice is then represented as a general "social contract" or agreement by free and equal persons on the basic terms of cooperation for their society. 2. The Original Position and Social Contract Doctrine Historically the idea of a social contract had a more limited role than Rawls assigns to it. In Thomas Hobbes and John Locke the social contract serves as an argument for the legitimacy of political authority. Hobbes argues that in a pre-social state of nature it would be rational for all to agree to authorize one person to exercise the absolute political power needed to maintain peace and enforce laws necessary for productive social cooperation. (Hobbes, 1651) By contrast, Locke argued against absolute monarchy by contending that no existing political constitution is legitimate or just unless it could be contracted into starting from a position of equal right with (a relatively peaceful) state of nature, and without violating any natural rights or duties. (Locke, 1690) For Rousseau and perhaps Kant too, the idea of a social contract plays a different role. It is an "idea of reason" (Kant) depicting a point of view that lawmakers and citizens should adopt in their reasoning to ascertain the "general will," which enables them to assess existing laws and decide upon measures that promote justice and citizens' common good. (Rousseau, 1762; Kant, 1793, 436-7; Kant 1797, 480) Rawls generalizes on Locke's, Rousseau's and Kant's natural rights theories of the social contract (TJ viii/viii rev.; 322/8 rev.). The purpose of the original position is to yield principles of justice that are fair to all persons in a society. The original position is a hypothetical perspective that we can adopt in our moral reasoning about the most basic principles of social and political justice. What primarily distinguishes Rawls's impartial perspective from its antecedents (in Hume, Smith, Kant, etc.) is that, rather than representing the judgment of one person, it is conceived socially, as a general agreement by (representatives of) all adult members of an ongoing society. The point of view of justice is then represented as a general "social contract" or agreement by free and equal persons on the basic terms of cooperation for their society. 2. The Original Position and Social Contract Doctrine Historically the idea of a social contract had a more limited role than Rawls assigns to it. In Thomas Hobbes and John Locke the social contract serves as an argument for the legitimacy of political authority. Hobbes argues that in a pre-social state of nature it would be rational for all to agree to authorize one person to exercise the absolute political power needed to maintain peace and enforce laws necessary for productive social cooperation. (Hobbes, 1651) By contrast, Locke argued against absolute monarchy by contending that no existing political constitution is legitimate or just unless it could be contracted into starting from a position of equal right with (a relatively peaceful) state of nature, and without violating any natural rights or duties. (Locke, 1690) For Rousseau and perhaps Kant too, the idea of a social contract plays a different role. It is an "idea of reason" (Kant) depicting a point of view that lawmakers and citizens should adopt in their reasoning to ascertain the "general will," which enables them to assess existing laws and decide upon measures that promote justice and citizens' common good. (Rousseau, 1762; Kant, 1793, 436-7; Kant 1797, 480) Rawls generalizes on Locke's, Rousseau's and Kant's natural rights theories of the social contract (TJ viii/viii rev.; 322/8 rev.). The purpose of the original position is to yield principles of justice that are fair to all persons in a society. The original position is a hypothetical perspective that we can adopt in our moral reasoning about the most basic principles of social and political justice. What primarily distinguishes Rawls's impartial perspective from its antecedents (in Hume, Smith, Kant, etc.) is that, rather than representing the judgment of one person, it is conceived socially, as a general agreement by (representatives of) all adult members of an ongoing society. The point of view of justice is then represented as a general "social contract" or agreement by free and equal persons on the basic terms of cooperation for their society. 2. The Original Position and Social Contract Doctrine Historically the idea of a social contract had a more limited role than Rawls assigns to it. In Thomas Hobbes and John Locke the social contract serves as an argument for the legitimacy of political authority. Hobbes argues that in a pre-social state of nature it would be rational for all to agree to authorize one person to exercise the absolute political power needed to maintain peace and enforce laws necessary for productive social cooperation. (Hobbes, 1651) By contrast, Locke argued against absolute monarchy by contending that no existing political constitution is legitimate or just unless it could be contracted into starting from a position of equal right with (a relatively peaceful) state of nature, and without violating any natural rights or duties. (Locke, 1690) For Rousseau and perhaps Kant too, the idea of a social contract plays a different role. It is an "idea of reason" (Kant) depicting a point of view that lawmakers and citizens should adopt in their reasoning to ascertain the "general will," which enables them to assess existing laws and decide upon measures that promote justice and citizens' common good. (Rousseau, 1762; Kant, 1793, 436-7; Kant 1797, 480) Rawls generalizes on Locke's, Rousseau's and Kant's natural rights theories of the social contract (TJ viii/viii rev.; 322/8 rev.). The purpose of the original position is to yield principles of justice that are fair to all persons in a society. The original position is a hypothetical perspective that we can adopt in our moral reasoning about the most basic principles of social and political justice. What primarily distinguishes Rawls's impartial perspective from its antecedents (in Hume, Smith, Kant, etc.) is that, rather than representing the judgment of one person, it is conceived socially, as a general agreement by (representatives of) all adult members of an ongoing society. The point of view of justice is then represented as a general "social contract" or agreement by free and equal persons on the basic terms of cooperation for their society. 2. The Original Position and Social Contract Doctrine Historically the idea of a social contract had a more limited role than Rawls assigns to it. In Thomas Hobbes and John Locke the social contract serves as an argument for the legitimacy of political authority. Hobbes argues that in a pre-social state of nature it would be rational for all to agree to authorize one person to exercise the absolute political power needed to maintain peace and enforce laws necessary for productive social cooperation. (Hobbes, 1651) By contrast, Locke argued against absolute monarchy by contending that no existing political constitution is legitimate or just unless it could be contracted into starting from a position of equal right with (a relatively peaceful) state of nature, and without violating any natural rights or duties. (Locke, 1690) For Rousseau and perhaps Kant too, the idea of a social contract plays a different role. It is an "idea of reason" (Kant) depicting a point of view that lawmakers and citizens should adopt in their reasoning to ascertain the "general will," which enables them to assess existing laws and decide upon measures that promote justice and citizens' common good. (Rousseau, 1762; Kant, 1793, 436-7; Kant 1797, 480) Rawls generalizes on Locke's, Rousseau's and Kant's natural rights theories of the social contract (TJ viii/viii rev.; 322/8 rev.). The purpose of the original position is to yield principles of justice that are fair to all persons in a society. The original position is a hypothetical perspective that we can adopt in our moral reasoning about the most basic principles of social and political justice. What primarily distinguishes Rawls's impartial perspective from its antecedents (in Hume, Smith, Kant, etc.) is that, rather than representing the judgment of one person, it is conceived socially, as a general agreement by (representatives of) all adult members of an ongoing society. The point of view of justice is then represented as a general "social contract" or agreement by free and equal persons on the basic terms of cooperation for their society. 2. The Original Position and Social Contract Doctrine Historically the idea of a social contract had a more limited role than Rawls assigns to it. In Thomas Hobbes and John Locke the social contract serves as an argument for the legitimacy of political authority. Hobbes argues that in a pre-social state of nature it would be rational for all to agree to authorize one person to exercise the absolute political power needed to maintain peace and enforce laws necessary for productive social cooperation. (Hobbes, 1651) By contrast, Locke argued against absolute monarchy by contending that no existing political constitution is legitimate or just unless it could be contracted into starting from a position of equal right with (a relatively peaceful) state of nature, and without violating any natural rights or duties. (Locke, 1690) For Rousseau and perhaps Kant too, the idea of a social contract plays a different role. It is an "idea of reason" (Kant) depicting a point of view that lawmakers and citizens should adopt in their reasoning to ascertain the "general will," which enables them to assess existing laws and decide upon measures that promote justice and citizens' common good. (Rousseau, 1762; Kant, 1793, 436-7; Kant 1797, 480) Rawls generalizes on Locke's, Rousseau's and Kant's natural rights theories of the social contract (TJ viii/viii rev.; 322/8 rev.). The purpose of the original position is to yield principles of justice that are fair to all persons in a society. The original position is a hypothetical perspective that we can adopt in our moral reasoning about the most basic principles of social and political justice. What primarily distinguishes Rawls's impartial perspective from its antecedents (in Hume, Smith, Kant, etc.) is that, rather than representing the judgment of one person, it is conceived socially, as a general agreement by (representatives of) all adult members of an ongoing society. The point of view of justice is then represented as a general "social contract" or agreement by free and equal persons on the basic terms of cooperation for their society. 2. The Original Position and Social Contract Doctrine Historically the idea of a social contract had a more limited role than Rawls assigns to it. In Thomas Hobbes and John Locke the social contract serves as an argument for the legitimacy of political authority. Hobbes argues that in a pre-social state of nature it would be rational for all to agree to authorize one person to exercise the absolute political power needed to maintain peace and enforce laws necessary for productive social cooperation. (Hobbes, 1651) By contrast, Locke argued against absolute monarchy by contending that no existing political constitution is legitimate or just unless it could be contracted into starting from a position of equal right with (a relatively peaceful) state of nature, and without violating any natural rights or duties. (Locke, 1690) For Rousseau and perhaps Kant too, the idea of a social contract plays a different role. It is an "idea of reason" (Kant) depicting a point of view that lawmakers and citizens should adopt in their reasoning to ascertain the "general will," which enables them to assess existing laws and decide upon measures that promote justice and citizens' common good. (Rousseau, 1762; Kant, 1793, 436-7; Kant 1797, 480) Rawls generalizes on Locke's, Rousseau's and Kant's natural rights theories of the social contract (TJ viii/viii rev.; 322/8 rev.). The purpose of the original position is to yield principles of justice that are fair to all persons in a society. The original position is a hypothetical perspective that we can adopt in our moral reasoning about the most basic principles of social and political justice. What primarily distinguishes Rawls's impartial perspective from its antecedents (in Hume, Smith, Kant, etc.) is that, rather than representing the judgment of one person, it is conceived socially, as a general agreement by (representatives of) all adult members of an ongoing society. The point of view of justice is then represented as a general "social contract" or agreement by free and equal persons on the basic terms of cooperation for their society. 2. The Original Position and Social Contract Doctrine Historically the idea of a social contract had a more limited role than Rawls assigns to it. In Thomas Hobbes and John Locke the social contract serves as an argument for the legitimacy of political authority. Hobbes argues that in a pre-social state of nature it would be rational for all to agree to authorize one person to exercise the absolute political power needed to maintain peace and enforce laws necessary for productive social cooperation. (Hobbes, 1651) By contrast, Locke argued against absolute monarchy by contending that no existing political constitution is legitimate or just unless it could be contracted into starting from a position of equal right with (a relatively peaceful) state of nature, and without violating any natural rights or duties. (Locke, 1690) For Rousseau and perhaps Kant too, the idea of a social contract plays a different role. It is an "idea of reason" (Kant) depicting a point of view that lawmakers and citizens should adopt in their reasoning to ascertain the "general will," which enables them to assess existing laws and decide upon measures that promote justice and citizens' common good. (Rousseau, 1762; Kant, 1793, 436-7; Kant 1797, 480) Rawls generalizes on Locke's, Rousseau's and Kant's natural rights theories of the social contract (TJ viii/viii rev.; 322/8 rev.). The purpose of the original position is to yield principles of justice that are fair to all persons in a society. The original position is a hypothetical perspective that we can adopt in our moral reasoning about the most basic principles of social and political justice. What primarily distinguishes Rawls's impartial perspective from its antecedents (in Hume, Smith, Kant, etc.) is that, rather than representing the judgment of one person, it is conceived socially, as a general agreement by (representatives of) all adult members of an ongoing society. The point of view of justice is then represented as a general "social contract" or agreement by free and equal persons on the basic terms of cooperation for their society. 2. The Original Position and Social Contract Doctrine Historically the idea of a social contract had a more limited role than Rawls assigns to it. In Thomas Hobbes and John Locke the social contract serves as an argument for the legitimacy of political authority. Hobbes argues that in a pre-social state of nature it would be rational for all to agree to authorize one person to exercise the absolute political power needed to maintain peace and enforce laws necessary for productive social cooperation. (Hobbes, 1651) By contrast, Locke argued against absolute monarchy by contending that no existing political constitution is legitimate or just unless it could be contracted into starting from a position of equal right with (a relatively peaceful) state of nature, and without violating any natural rights or duties. (Locke, 1690) For Rousseau and perhaps Kant too, the idea of a social contract plays a different role. It is an "idea of reason" (Kant) depicting a point of view that lawmakers and citizens should adopt in their reasoning to ascertain the "general will," which enables them to assess existing laws and decide upon measures that promote justice and citizens' common good. (Rousseau, 1762; Kant, 1793, 436-7; Kant 1797, 480) Rawls generalizes on Locke's, Rousseau's and Kant's natural rights theories of the social contract (TJ viii/viii rev.; 322/8 rev.). The purpose of the original position is to yield principles of justice that are fair to all persons in a society. The original position is a hypothetical perspective that we can adopt in our moral reasoning about the most basic principles of social and political justice. What



Rawls contends that a conception of justice should enable citizens to fully exercise and adequately develop their moral powers, including their capacity for justice. It should then engage their sense of justice in such a way that they can regard justice as a burden that should be accepted as an end in itself, not as a means to other ends. Rawls argues that a conception of justice should be based on an understanding of the reasonable comprehensive philosophical, religious, or moral doctrines that reasonable persons affirm (as Rawls contends later in Political Liberalism). For then justice and the full and informed exercise of the sense of justice are for reasonable and rational persons essential goods, preconditions for their living a good life, as that is defined by their rational conception of the good. 6. The Arguments for the Principles of Justice from the Original Position The original position is not a bargaining situation where the parties make proposals and counterproposals and negotiate over different principles of justice. Nor is it a wide ranging discussion where the parties debate, deliberate, and design their own conception of justice (unlike, for example, Habermas's discourse ethics; see Habermas, 1995). Instead, the parties' deliberations are much more constrained and regulated. They are presented with a list of conceptions of justice taken from the tradition of western political philosophy. These include different versions of utilitarianism, perfectionism, and intuitionism (or pluralist views), rational egoism, justice as fairness, and a group of "mixed conceptions" that combine elements of these. (For Rawls's initial list see TJ 124/107) Rawls later says libertarian entitlement principles should also be added to the list, and contends the principles of justice are still preferable. (JF 83). (Nozick agrees and says the OP is incapable of yielding historical entitlement principles, but only patterned end-state principles instead. Nozick 1974, 198-204. Rawls replies that the difference principle does not conform to any observable pattern but grounds fair distributions in a fair social process that must actually be carried out. PL 282-83) The parties' deliberations are confined to discussing and agreeing upon the conception that each finds most rational, given their specified interests. In a series of pairwise comparisons they consider all the conceptions of justice made available to them and ultimately agree unanimously to accept the conception that survives this winnowing process. In this regard, the original position is best seen as a kind of selection process wherein the parties' deliberations are constrained by the background conditions imposed by the original position as well as the list of conceptions of justice provided to them. They are assigned the task of agreeing on principles for designing the basic structure of a self-contained society under the circumstances of justice. In making their decision, the parties are motivated only by their own rational interests. They do not take moral considerations of justice into account except in so far as these considerations bear on their achieving their interests within society. Their interests again are defined in terms of their each acquiring an adequate share of primary social goods (rights and liberties, powers and opportunities, income and wealth, etc.) and achieving the background social conditions enabling them to effectively pursue their conception of the good and realize their higher-order interests in developing and exercising their moral powers. Since the parties are ignorant of their particular conceptions of the good and of all other particular facts about their society, they are not in a position to engage in bargaining. In effect they all have the same general information and are motivated by the same interests. Rawls makes four arguments in Theory, Part I for the principles of justice. The main argument for the difference principle is made later in TJ §49, and is amended and clarified in Justice as Fairness: A Restatement. The common theme throughout the original position arguments is that it is more rational for the parties to choose the principles of justice over any other alternative. Rawls devotes most of his attention to the comparison of justice as fairness with classical and average utilitarianism, with briefer discussions of perfectionism (TJ, §50) and intuitionism (TJ 278-81) Here I'll focus discussion primarily on Rawls's comparison between justice as fairness and utilitarianism. 6.1 The Principles of Justice Before turning to Rawls's arguments from the original position, it is helpful to have available the principles of justice and other principles that constitute Justice as Fairness. First Principle: "Each person has an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all." (TJ 266) The first principle was revised in 1982 to say "Each person has an equal right to a fully adequate scheme of equal basic liberties. . ." (PL 291) replacing "... the most extensive scheme of equal basic liberties." Notably, Rawls also introduces in Political Liberalism, almost in passing, a principle of basic needs that precedes the first principle and requires that citizens' basic needs be met at least to the extent that they can understand and fruitfully exercise their basic rights and liberties. (PL 7, JF 79n.) This social minimum is also said in Political Liberalism to be a "constitutional essential" for any reasonable liberal conception of justice. (PL 166, 228ff.; JF 47, n.7) The basic rights and liberties protected by the first principle are specified by a list (see TJ 53f., PL 291): liberty of conscience and freedom of association, (TJ §83-4); freedom of thought and freedom of speech and expression (PL pp.340-363); the integrity and freedom of the person and the right to hold personal property; equal rights of political participation and their fair value (TJ §536-37); and the rights and liberties protected by the rule of law (due process, freedom from arbitrary arrest, etc. TJ §38). (Rawls says the right to ownership of means of production and laissez faire freedom of contract are not included among the basic liberties. TJ, 54 rev. Also freedom of movement and free choice of occupation are said to be primary goods protected by fair equality of opportunity principle. PL 76, JF 58f.) Second Principle: "Social and economic inequalities are to satisfy two conditions. First they must attach to offices and positions open to all under conditions of fair equality of opportunity; and second they must be to the greatest advantage of the least advantaged members of society [the difference principle]" consistent with the just savings principle. (PL 281, JF 42-43, TJ 301/266 rev.) Just Savings Principle: Each generation should save for future generations at a savings rate that they could rationally expect past generations to have saved for them. (TJ §44; JF 159-160) Principles for individuals, include (a) the natural duties to uphold justice, mutual respect, mutual aid, and not to injure or harm the innocent (TJ §819, 51); and (b) the principle of fairness, to do one's fair share in just or nearly just practices and institutions from which one accepts their benefits, (which grounds the principle of reciprocity, to keep one's promises and commitments. (TJ §818, 52). The Priority Principles, the principles of justice are ranked in lexical order. (a) The priority, or liberty requires that basic liberties can only be restricted to strengthen the system of liberties shared by all. (b) Fair equality of opportunity is lexically prior to the difference principle. (c) The second principle is prior to the principle of efficiency and maximizing the sum of advantages. (TJ 302/266 rev.) The General Conception of Justice: "All social 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 1971, 302. Note: The general conception is the difference principle generalized to all primary goods (TJ 1971, 83); it applies in non-ideal conditions where the priority of liberty and opportunity is not sustainable. 6.2 The Argument from the Maximin Criterion (TJ, §826-28) Describing the parties' choice as a rational choice subject to the reasonable constraints imposed by the original position allows Rawls to invoke the theory of rational choice and decision under conditions of uncertainty. In rational choice theory there are a number of potential "strategies" or rules of choice that are more or less reliably used depending on the circumstances. One rule of choice—called "maximin"—directs that we play it as safe as possible by choosing the alternative whose worst outcome leaves us better off than the worst outcome of all other alternatives. The aim is to "maximize the minimum" regret or loss to one's position (measured in terms of welfare or, for Rawls, one's share of primary social goods). To follow this strategy, Rawls says you should choose as if your enemy were to assign your social position in whatever kind of society you end up in. By contrast another strategy leads us to focus on the most advantaged position and says we should "maximize the maximum" potential gain—"maximax"—and choose the alternative whose best outcome leaves us better off than all other alternatives. Which, if either, of these strategies is more sensible to use depends on the circumstances and many other factors. A third strategy advocated by orthodox Bayesian decision theory, says we should always choose to directly maximize expected utility. To do so under conditions of uncertainty of outcomes, the degree of uncertainty should be factored into one's utility function, with probability estimates assigned to alternatives based on the limited knowledge that one has. Given these subjective estimates of probability incorporated into one's utility function, one can always choose the alternative that maximizes expected utility. Since it simplifies matters to apply the same rule of choice to all decisions this is a highly attractive idea, so long as one can accept that it is normally safe to assume that the maximization of expected utility leads over time to maximizing actual utility. What about those extremely rare instances where there is absolutely no basis upon which to make probability estimates? Suppose you don't even have a hunch regarding the greater likelihood of one alternative over another. According to orthodox Bayesian decision theory, the "principle of insufficient reason" should then be observed; it says that when there is no reason to assign a greater likelihood to one alternative rather than another, then an equal probability is to be assigned to each potential outcome. This makes sense on the assumption that if you have no more premonition of the likelihood of one option rather than another, they are for all you know equally likely to occur. By observing this rule of choice consistently over time, a rational chooser presumably should maximize expected individual utility, and hopefully actual utility as well. What now is the appropriate decision rule to be used to choose principles of justice under conditions of complete uncertainty of probabilities in Rawls's original position? Rawls argues that, given the enormous gravity of choice in the original position, plus the fact that the choice is not repeatable (there's no opportunity to renegotiate or revise one's decision), it is rational for the parties to follow the maximin strategy when choosing between the principles of justice and principles of average or aggregate utility (or any other principles that do not guarantee basic rights, liberties, opportunities, and a social minimum). Not surprisingly, following the maximin rule of choice results in choice of the principles of justice over the principles of utility (average or aggregate); for unlike utilitarianism, justice as fairness guarantees equal basic liberties, fair equal opportunities, and an adequate social minimum for all citizens. Why does Rawls think maximin is the rational choice rule? Recall what is at stake in choice from the original position. The decision is not an ordinary choice. It is rather a unique and irrevocable choice where the parties decide the basic structure of their society, or the kind of social world they will live in and the background condition of their position, and there is no hope of renegotiating the outcome, a rational person would agree to the principles of justice instead of the principle of utility. For the principles of justice imply that no matter what position you occupy in society, you will have the rights and resources needed to maintain your valued commitments and purposes, to effectively exercise your capacities for rational and moral deliberation and action, and to maintain your sense of self-respect as an equal citizen. With the principle of utility there is no such guarantee; everything is "up for grabs" (so to speak) and subject to loss if required by the greater sum of utilities. Conditions (2) and (3) for applying maximin are then satisfied in the comparison of justice as fairness with the principle of (average or aggregate) utility. It is often claimed that Rawls's parties are "risk-averse;" otherwise they would never follow the maximin rule but would take a chance on riskier but more rewarding outcomes provided by the principle of utility. Thus, John Harsanyi contends that it is more rational under conditions of complete uncertainty always to choose according to the principle of insufficient reason and assume an equal probability of occupying any position in society. When the equiprobability assumption is made, the parties in the original position would choose the principle of average utility instead of the principles of justice (Harsanyi 1975). Rawls denies that the parties have a psychological disposition to risk-aversion. They have no knowledge of their attitudes towards risk. He argues however that it is rational to choose as if one were risk averse under the highly exceptional circumstances of the original position. His point is that, while there is nothing rational about a fixed disposition to risk in general, it is nonetheless rational in some circumstances to choose conservatively to protect certain fundamental interests against loss or compromise. It does not make one a risk averse person, but instead it is normally rational to purchase auto liability, health, and home insurance to guard against accident or calamity (assuming it is affordable). The original position is such a situation writ large. Even if one knew in the original position that citizen one represents enjoys gambling and taking great risks, this would still not be a reason to gamble with their rights, liberties and starting position in society. For if the high risktaker were born into a traditional, repressive, or fundamentalist society, they might never have an opportunity for gambling and taking other risks they normally enjoy. It is rational then even for high risktakers to choose conservatively in the original position and guarantee their future opportunities to gamble or otherwise take risks. Harsanyi and other orthodox Bayesians contend that maximin is an irrational decision rule, and they provide ample examples. To take Rawls' own example, in a lottery where the loss and gain alternatives are either (0, n) or (1/n, 1) for all natural numbers n, maximin says choose the latter alternative (1/n, 1). This is clearly irrational for almost any number n except very small numbers. (TJ 136 rev.). But such examples do not suffice here; simply because maximin is under most circumstances irrational does not mean that it is never rational. For example, suppose n>1 and you must have 1/n to save you own life. Given the gravity of the circumstances, it would be rational to choose conservatively since you are guaranteed 1/n according to the maximin strategy, and there is no guarantee you will survive if you choose according to the principle of insufficient reason. No doubt maximin is an irrational strategy under most circumstances of choice uncertainty, particularly under circumstances where we will have future opportunities to recoup our losses and choose again. But these are not the circumstances of the original position. Once the principles of justice are decided, they apply in perpetuity, and there is no opportunity to renegotiate or escape the situation. One who relies on the equiprobability assumption in choosing principles of justice in the original position is being foolishly reckless given the gravity of choice at stake. It is not being risk-averse, but rather entirely rational to refuse to gamble with one's basic liberties, fair equal opportunities and adequate resources needed to pursue one's most cherished ends and commitments, simply for the unknown chance of gaining the marginally greater social powers, income and wealth that might be available to some in a society governed entirely by the principle of utility. Rawls exhibits the force of the maximin argument in discussing liberty of conscience. He says (TJ, sect. 33) that a person who is willing to jeopardize their right to hold and practice their conscientious religious, philosophical and moral convictions, all for the sake of gaining uncertain added benefits via the principle of utility, does not know what it means to have conscientious beliefs, or at least does not take such beliefs seriously (TJ 207-08/181-82 rev.). A rational person with convictions about what gives life meaning is not willing to negotiate with and gamble away the right to hold and express those convictions and the freedom to act on them. After all what could be the basis for negotiation, for what could matter more than the objects of one's most sincere convictions and commitments? Some people (e.g. some nihilists) may not have any conscientious convictions (except the belief that nothing is worthwhile) and are simply willing to act on impulse or on whatever thoughts and desires they are unwilling to compromise. (Besides, even the nihilist should want to protect the freedom to be a nihilist, to avoid ending up in an intolerant religious society.) Thus it remains irrational to jeopardize basic liberties by choosing the principle of utility instead of the principles of justice. None of this is to say that maximin is normally a rational choice strategy. Rawls himself says it "is not, in general, a suitable guide for choices under uncertainty" (TJ 153). It is not even a rational strategy in the original position when the alternatives for choice guarantee basic liberties, equal opportunities, and a social minimum guaranteed by the principle of average utility - see the discussion in the supplementary section: The Argument for the Difference Principle in the supplementary document The Argument for the Difference Principle and the Four Stage Sequence. Rawls relies upon the maximin argument mainly to argue for the first principle of justice and fair equality of opportunity. Other arguments are needed to support his claim that justice requires the social minimum be determined by the difference principle. 6.3 The Strains of Commitment There are three additional arguments Rawls makes to support justice as fairness (all in TJ, sect. 29). Each of these depends upon the concept of a "well-ordered society." The parties in the original position are to choose principles that are to govern a well-ordered society where everyone agrees, complies with, and wants to comply with its principles of justice. The ideal of a well-ordered society is Rawls's development of social contract doctrine. It is a society in which (1) everyone knows and willingly accepts and affirms the same public principles of justice and everyone knows this; (2) these principles are successfully realized in basic social institutions, including laws and conventions, and are generally complied with by citizens; and (3) reasonable persons are morally motivated to comply by their sense of justice - they want to do what justice requires of them (TJ 4-5, §69). There are then two sides to Rawls's social contract. The parties in the original position have the task of agreeing to principles that all can rationally accept behind the veil of ignorance under the conditions of the original position. But their rational choice is partially determined by the principles that free and equal moral persons in a well ordered society who are motivated by their sense of justice reasonably can accept, agree to and comply with, as the basic principles governing their social and political relations. The parties are to assess principles according to the relative stability of the well ordered societies into which they are incorporated. Thus a well-ordered society of justice as fairness is to be compared with a well-ordered society whose basic structure is organized according to the average utility principle, aggregate utility, perfectionism, intuitionism, libertarianism, and so on. They are to consider which of these societies' basic struture is relatively more stable and likely to endure over time from one generation to the next, given natural and socially influenced psychological propensities and conditions of social cooperation as they interact with alternative principles of justice. Now to return to Rawls's arguments for his principles of justice. The first of Rawls's three arguments highlights the idea that choice in the original position is an agreement, and involves certain "strains of commitment." It is assumed by all the parties that all will comply with the principles they agree to once the veil is lifted and they are members of a well-ordered society (TJ 176f./153f. and CP 250ff). Knowing that they will be held to their commitment and expected to comply with principles for a well-ordered society, the parties must choose principles that they sincerely believe they will be able to accept, endorse and willingly observe under conditions where these principles are generally accepted and enforced. For reasons to be discussed shortly, Rawls says this condition favors agreement on the principles of justice over utilitarianism and other alternatives. But first, consider the frequent objection that there is no genuine agreement in the original position, for the thick veil of ignorance deprives the parties of all bases for bargaining (cf. TJ, 139-40/120-21 rev.). In the absence of bargaining, it is said, there can be no contract. For contracts must involve a quid pro quo—something given for something received (called "consideration" at common law). The parties in the OP cannot bargain without knowing what they have to offer or to gain in exchange. So (the objection continues) Rawls's original position does not involve a real social contract, unlike those that transpire, say, in a state of nature. Rather, since the parties are all "described in the same way," there is no need for multiple parties but simply the rational choice of one person in the original position (see Hampton, 1980, 334; see also Gauthier, 1974 and 1985, 203). In response, not all contracts involve bargaining or are of the nature of economic transactions. Some involve a mutual pledge and commitment to shared purposes and principles. Marriage contracts, or agreements among friends or the members of a religious, benevolent, or political association are often of this nature. For example, the Mayflower Compact was a "covenant" to "combine ourselves together into a civil body politic" charged with making and administering "just and equal laws." For the general good. Likewise the U.S. Constitution represents itself as a commitment wherein "We The People ... do ordain and establish this Constitution" in order "to establish justice," "promote the general welfare," "secure the blessings of liberty," and so on. The agreement in Rawls's original position is more of this nature. Even though ignorant of particular facts about themselves, the parties in fact do give something in exchange for something received: they all exchange their mutual commitment to accept and abide by the principles of justice and to uphold just institutions once they enter their well-ordered society. Each agrees only on condition others do too, and all tie themselves into social and political relations in perpetuity. Their agreement is final, and they will not permit its renegotiation should circumstances turn out to be different than some had hoped for. Their mutual commitment to justice is reflected by the fact that once these principles become embodied in institutions there are no legitimate means that permit anyone to depart from the terms of their agreement. As a result, the parties have to take seriously the moral and legal obligations and potential social sanctions they will incur as a result of their agreement, for there is no going back to the initial situation. So if they do not sincerely believe that they can accept the requirements of a conception of justice and voluntarily conform their actions and life plans accordingly, then these are strong reasons to avoid choosing those principles. It would not be rational for the parties to take risks, falsely assuming that if they end up badly, they can violate at will the terms of agreement or later regain their initial situation and renegotiate terms of cooperation (see Freeman, 1990; Freeman, 2007b, 180-182). Rawls gives special poignancy to this mutual commitment of the parties by making it a condition that the parties cannot choose and agree to principles in bad faith; they have to be able, not simply to live with and grudgingly accept, but instead to willingly endorse the principles of justice as members of society. Essential to Rawls's argument for stability is the assumption of everyone's willing compliance with requirements of justice. This is a feature of a well-ordered society. The parties are assumed to have a sense of justice; indeed the development and exercise of it is one of their fundamental interests. Hence they must choose principles that that they can not only accept and live with, but which are responsive to their sense of justice and they can willingly endorse. Given these conditions on choice, the parties cannot take risks with principles they know they will have difficulty complying with voluntarily. They would be making an agreement in bad faith, and this is ruled out by the conditions of the original position. Rawls contends that these "strains of commitment" created by the parties' agreement strongly favor the principles of justice over the principles of utility and other teleological (and most consequentialist) views. For everyone's freedom, basic rights and liberties, and basic needs are met by the principles of justice because of their egalitarian nature. Given the lack of these guarantees under the principle of utility, it is much more difficult for those who end up worse off in a utilitarian society to willingly accept their situation and commit themselves to the utility principle. It is a rare person indeed who can freely and without resentment sacrifice their life prospects so that those who are better off can have even greater comforts, privileges, and powers. This is too much to demand of our capacities for human benevolence. It requires a kind of commitment that people cannot make in good faith, for who could willingly support laws that are so detrimental to oneself and the people one cares about most that they must sacrifice their fundamental interests for the sake of those more advantaged? Besides, why should we encourage such subservient dispositions and the accompanying lack of self-respect? The principles of justice, by contrast, conform better with everyone's interests, their desire for self-respect and their natural moral capacities to reciprocally recognize and respect others' legitimate interests while freely promoting their own good. The strains of commitment incurred by agreement in the original position provide strong reasons for the parties to choose the principles of justice and reject the risks involved in choosing the principles of average or aggregate utility. 6.4 Stability, Publicity, and Self-Respect Rawls's strains-of-commitment argument explicitly relies upon a rarely noted feature of his argument, as mentioned earlier, there are in effect two social contracts. First, hypothetical agents situated equally in the original position unanimously agree to principles of justice. This agreement has attracted the most attention from Rawls's critics. But the parties' hypothetical agreement in the original position is patterned on the general acceptability of a conception of justice by free and equal persons in a well-ordered society, Rawls says, "The reason for invoking the concept of a contract in the original position lies in its correspondence with the features of a well-ordered society [which] require... that everyone accepts, and knows that the others accept, the same principles of justice" (CP 250). In order for the hypothetical parties in the original position to agree on principles of justice, there must be a high likelihood that real persons, given human nature and general facts about social and economic cooperation, can also agree and act on the same principles, and that a society structured by these principles is feasible and can endure. This is the stability requirement referred to earlier. One conception of justice is relatively more stable than another the more willing people are to observe its requirements under conditions of a well-ordered society. Assuming that each conception of justice has a corresponding society that is as well-ordered as can be according to its terms, the stability question raised in Theory is: Which conception of justice is more likely to engage the moral sensibilities and sense of justice of free and equal persons as well as affirm their good? This requires an inquiry into moral psychology and the human good, which takes up most of Part III of A Theory of Justice. Rawls makes two arguments in Theory from the original position that invoke the stability requirement, the arguments (1) from publicity and (2) from self-respect (see TJ, §29) (1) The argument from publicity: Rawls contends that utilitarianism, perfectionism, and other "teleological" conceptions are unlikely to be freely acceptable to many citizens when made fully public under the conditions of a well-ordered society. Recall the publicity condition discussed earlier: A feature of a well-ordered society is that its regulative principles of justice are publicly known and appealed to as a basis for deciding laws and justifying basic institutions. Because all reasonable members of society accept the public conception of justice, there is no need for the illusions and delusions of ideology for a society to function properly and citizens to accept its laws and institutions willingly. In this sense a well-ordered society lacks false consciousness about the bases of social and political relations. (PL 68-69n.) A conception of justice that satisfies the publicity condition but that cannot maintain the stability of a well-ordered society is to be rejected by the parties in the original position. Rawls contends that under the publicity condition justice as fairness generally engages citizens' sense of justice and remains more stable than utilitarianism (TJ 177f./154f rev.) For public knowledge that reasons of maximum average (or aggregate) utility determine the distribution of benefits and burdens would lead those worse-off to object to and resent their situation, and reject the principle of utility as the basic principle governing social institutions. After all, the well-being and interests of the least advantaged, perhaps even their basic liberties, are being sacrificed for the greater happiness of those who are already more fortunate and have a greater share of primary social goods. It is too much to expect of human nature that people should freely acquiesce in and embrace such publicly known terms of cooperation. By contrast, the principles of justice are designed to advance reciprocally everyone's position; those who are better off do not achieve their gains at the expense of the less advantaged. "Since everyone's good is affirmed, all acquire inclinations to uphold the scheme" (TJ, 177/155). It is a feature of our moral psychology, Rawls contends, that we normally come to form attachments to people and institutions that are concerned with our good; moreover we tend to resent those persons and institutions that take unfair advantage of us and act contrary to our good. Rawls argues at length in chapter 8 of Theory, §570-75, that justice as fairness accords better than alternative principles with the reciprocity principles of moral psychology that are characteristic of human beings' moral development. In Political Liberalism, Rawls expands the publicity condition to include three levels: First, the principles of justice governing a well-ordered society are publicly known and appealed to in political debate and deliberation; second, so too are the general beliefs in light of which society's conception of justice is generally accepted—including beliefs about human nature and the way political and social institutions generally work—and citizens generally agree on these beliefs that support society's conception of justice. Finally the full justification of the public conception of justice is also publicly known (or at least publicly available to any who are interested) and is reflected in society's system of law, judicial decisions and other political institutions, as well as its system of education. (2) The argument from the social bases of self-respect: The publicity condition is also crucial to Rawls's fourth argument for the principles of justice, from the social bases of self-respect (TJ, 178-82/155-59 rev.). These principles, when publicly known, give greater support to citizens' sense of self-respect than do utilitarian and perfectionist principles. Rawls says self-respect is "perhaps the most important primary good," (TJ, 440/386 rev.) since few things seem worth doing if a person has little sense of their own worth or no confidence in their abilities to execute a worthwhile life plan or fulfill the duties and expectations in their role as citizens. The parties in the original position will then aim to choose principles that best secure their sense of self-respect. Now being regarded by others as a free and independent person of equal status with others is crucial to the self-respect of persons who regard themselves as free and equal members of a democratic society. Justice as fairness, by affording and protecting the priority of equal basic liberties and fair equal opportunities for all, secures the status of each as free and equal citizens. For example, because of equal political liberties, there are no "passive citizens" who must depend on others to politically protect their rights and interests; and with fair equal opportunities no one has grounds to experience the resentment that inevitably arises in societies where social positions are effectively closed off to those less advantaged or less powerful. Moreover, the second principle secures adequate social powers and economic resources for all so that they find the effective exercise of their equal basic liberties to be worthwhile. The second principle has the effect of making citizens socially and economically independent, so that no one need be subservient to the will of another. Citizens then can regard and respect one another as equals, and not as masters or subordinates. ("Non-domination," an idea central to contemporary Republicanism, is then essential to citizens' sense of self-respect in Rawls's sense. See Pettit 1997.) Equal basic liberties, fair equal opportunities, and political and economic independence are among the social bases of self-respect in a democratic society. The parties in the original position should then choose the principles of justice over utilitarianism and other teleological views both to secure their sense of self-respect, and to procure the same for others, thereby guaranteeing greater overall stability. In connection with Rawls's argument for the greater stability of principles of justice on grounds of their publicity and the bases of self-respect, Rawls provides a Kantian interpretation of difference principle. He says: "[T]he difference principle interprets the distinction between treating men as means only and treating them as ends in themselves. To regard persons as ends in themselves in the basic design of society is to agree to forgo those gains that do not contribute to everyone's expectations. By contrast to regard persons as means is to be prepared to impose on those already less favored still lower prospects of life for the sake of the higher expectations of others" (TJ 157 rev.). Rawls says the principle of utility does just this; it treats the less fortunate as means since it requires them to accept even lower life prospects for the sake of others who are more fortunate and already better off. This exhibits a lack of respect for the less advantaged and in turn has the effect of undermining their sense of self respect. (TJ 158 rev.) The difference principle, by contrast, does not treat people as means or undermine their sense of self respect, and this adds to the reasons the parties have for choosing the principles of justice instead of the principle of utility. Rawls substantially relies on the publicity condition to argue against utilitarianism and perfectionism. He says publicity "arises naturally from a contractarian standpoint" (TJ, 133/115 rev.). In Theory he puts great weight on publicity ultimately because he thinks that giving people knowledge of the moral bases of coercive laws and the principles governing society is a condition of fully acknowledging and respecting them as free and responsible rational moral agents. With publicity of principles of justice, people have knowledge of the real reasons for their social and political relations and the formative influences of the basic structure on their characters, plans and prospects. In a well-ordered society with a public conception of justice, there is no need for an "esoteric morality" that must be confined "to an enlightened few" (as Sidgwick says of utilitarianism, Sidgwick 1907 [1901], 490). Moreover, public principles of justice can serve agents in their practical reasoning and provide democratic citizens a common basis for political argument and justification. These considerations underlie Rawls's later contention that having knowledge of the principles that determine the bases of social relations is a precondition of individuals' freedom. (CP 325f.) Rawls means in part that publicity of society's fundamental principles is a condition of citizens' exercise of the powers and abilities that enable them to take full responsibility for their lives. Full publicity is then a condition of the political and (in TJ) moral autonomy of persons, which are significant values according to justice as fairness. (TJ §78, PL 68, CP 325-26) Utilitarians often regard Rawls's emphasis on the publicity of the fundamental principles underlying social cooperation as unwarranted. They contend that publicity of laws is of course important for them to be effective, but there's no practical need for the publicity of the fundamental principles (such as the principles of efficiency and utility) that govern political decisions, the economy, and society, much less so for the publicity of the full justification of these principles. Most people are not interested and have little understanding of the complex often technical details that must go into deciding laws and social policies. Moreover, as Sidgwick claimed, utilitarianism functions better as an "esoteric morality" that is not generally incorporated into the public justification of laws and institutions. Others claim that Rawls's arguments from publicity are exaggerated. If people were properly educated to believe that promoting greater overall happiness or welfare is the ultimate requirement of justice and more generally of morality, then just as they have for centuries constrained their conduct and their self-interests and accepted political constraints on their own liberties for the sake of their religious beliefs, so too could they be educated to accept the promotion of social utility and the general welfare as the fundamental bases for social and political cooperation. Supplementary Documents on Other Topics Additional topics concerning the original position are discussed in the following supplementary documents: The Argument for the Difference Principle. Explains the Difference Principle and the least advantaged class. Comparison of the difference principle with mixed conceptions, including restricted utility. Arguments from reciprocity, stability of self-respect, and the strains of commitment. Rawls's reasons why the difference principle supports property owning democracy rather than welfare-state capitalism. The Four Stage Sequence. How principles chosen in OP (first stage) apply to choice of political constitution (second-stage), democratic legislation (third stage), and application of laws to particular circumstances (fourth stage). Ideal Theory, Strict Compliance and the Well-Ordered Society. Why strict compliance is said to be necessary to justification of universal principles. Sen's, Mills's, and others' criticisms of ideal theory. Rawls's contention that ideal theory is necessary to determine injustice in non-ideal conditions. Role of non-ideal theory. A Liberal Feminist Critique of the Original Position and Justice within the Family. Criticism of "heads of families" assumption in OP and Rawls's response to criticisms that principles do not secure equal justice for women and children. Rawls's discussion of justice within the family. The Original Position and the Law of Peoples. Rawls's extension of OP to decisions on the Law of Peoples governing relations among liberal and decent societies. Human rights, the duty to assist burdened peoples, outlaw societies, and Rawls's rejection of a global principle of distributive justice. Constructivism, Objectivity, Autonomy, and the Original Position. Kantian Interpretation of the OP and Constructivism. OP as a procedure of construction and objective point of view. Response to Humean argument that social agreements cannot justify. Role of OP in reflective equilibrium. Is the Original Position Necessary or Relevant? Reply to claims that OP is superfluous or irrelevant. Why Rawls thinks rational acceptance of principles in OP and congruence of Right and Good is essential to justice.