

Journal of Political Science

Volume 9
Number 1 (*Fall*)

Article 3

November 1981

The Justice of Rawls' Original Position

Steven D. Ealy

Follow this and additional works at: <https://digitalcommons.coastal.edu/jops>



Part of the [Political Science Commons](#)

Recommended Citation

Ealy, Steven D. (1981) "The Justice of Rawls' Original Position," *Journal of Political Science*: Vol. 9 : No. 1 , Article 3.

Available at: <https://digitalcommons.coastal.edu/jops/vol9/iss1/3>

This Article is brought to you for free and open access by the Politics at CCU Digital Commons. It has been accepted for inclusion in Journal of Political Science by an authorized editor of CCU Digital Commons. For more information, please contact commons@coastal.edu.

The Justice of Rawls' Original Position

STEVEN D. EALY

Western Carolina University

I

According to John Rawls, a theory of justice is a description of "our sense of justice."¹ What does Rawls mean by "sense of justice"? What is it, exactly, that a theory of justice is to describe? Our sense of justice is "a normally effective desire to apply and to act upon the principles of justice, at least to a certain minimum degree" (505).

It may be, then, that the description of our sense of justice will be a psychological description, explaining to us the origins of this desire. But the psychological explanation of this sense appears to be something that Rawls seeks to overcome, rather than to use as a description. At least this seems to be the case for a "well-ordered society" (514).²

The description is not, then, psychological in nature, and neither is it a mere cataloging of our currently accepted judgments and the reasons that we adduce in their support. "Rather, what is required is a formulation of a set of principles which, when conjoined to our beliefs and knowledge of the circumstances, would lead us to make these judgments with their supporting reasons were we to apply these principles conscientiously and intelligently" (46).

The function of a theory of justice, then, appears to be to make explicit the already implicit kernel of all our judgments. While our actions would be no different after we knew the hidden principles upon which we have been acting all along, our "moral sensibility" would be different, for it would now be characterized by "a conception of justice" (46).

Rawls, however, is not interested in determining the principles that form the hidden nucleus of our everyday, real-life, actions; that would explain only our everyday, real-life, sense of justice. "From the standpoint of moral philosophy," Rawls argues,

¹ John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971), p. 46. All unidentified numbers within this paper refer to the page numbers in this book. Although Rawls indicates that this understanding of what a theory of justice is provisional, he does not revise that understanding later in the book.

² Rawls does present a three-stage "psychology of moral development" (462-79).

the best account of a person's sense of justice is not the one which fits his judgments prior to his examining any conception of justice, but rather the one which matches his judgments in reflective equilibrium (48).

While there are various interpretations of "reflective equilibrium," questions of moral philosophy are concerned with the case in which "one is to be presented with all possible descriptions to which one might plausibly conform one's judgments together with all relevant philosophical arguments for them" (49). Once one is confronted with these descriptions and arguments and has examined them, has matched principles with descriptions, descriptions with arguments, and arguments with principles—once one has done all of this and has separated the wheat from the chaff, the sheep from the goats, and the wise virgins from the foolish—so that one's "principles and judgments coincide . . . and [one] know[s] to what principles [one's] judgments conform and the premises of their derivation" (20), then one has arrived at "reflective equilibrium" (49).

Rawls is reasonable, however, and he does admit that "it is doubtful whether one can ever reach this state" (49; cf. 21). Because we cannot achieve reflect equilibrium by direct assault, Rawls decides to use a flanking maneuver and turning movement to arrive at the principles necessary for the description of our moral sense from the rear. To accomplish this task Rawls introduces a particular form of the social contract (11) which he calls "the original position" (12-14, 118ff.). According to Rawls, "the original position is a purely hypothetical situation." He continues: "The conception of the original position is not intended to explain human conduct except insofar as it tries to account for our moral judgments and helps to explain our having a sense of justice" (120). The principles of justice necessary for the description of our sense of justice will be derived from this original position (17-22, 118-26).

II

The original position is "a purely hypothetical situation characterized so as to lead to a certain conception of justice" (12). In the original position certain persons will select the principles of justice which will apply to the foundations of society. This position is characterized by two main features: the rationality of the parties involved (142ff) and the veil of ignorance (136-42).

The individuals in the original position are rational, that is, they will attempt to acknowledge principles which advance their "system of

ends as far as possible" (144). These persons define their system of ends in accordance with Rawls' account of "primary social goods" (142). Primary goods are "things that every rational man is presumed to want. These goods normally have a use whatever a person's rational plan of life" (62).³ Among primary social goods would be "rights and liberties, powers and opportunity, income and wealth" (62). The rationality of persons in the original position is thus defined by their preference for more primary social goods rather than less.

These rational persons will meet together and select, from a list of possible alternative principles, those principles of justice which "are to apply to the basic structure of society" (118-26, 250). Since each of these people is going to attempt to take as much of primary social goods for himself as possible, there is the danger that each person will vote on the principles of justice out of sheer self-interest. This possibility creates a problem for Rawls (18-19).

In the original position all are guided by self-interest, but all do not necessarily have the same abilities. Those with greater abilities or with better social position will be better able to secure those goods that all desire. This aspect of the social world, however, "seems arbitrary from a moral point of view" (15). It is to overcome this arbitrariness that the veil of ignorance is introduced. The veil of ignorance "ensures that no one is advantaged or disadvantaged in the choice of principles by the outcome of natural chance or the contingency of social circumstances" (12). The principles chosen from behind the veil of ignorance will be fair, because no one will be able to design principles to favor his particular condition.

Behind the veil of ignorance men know general facts about human society, but certain specific facts are hidden. No one in the original position knows his place in society, his social class, or his economic standing. While all of the people in the original position belong to the same generation, no one knows to which generation they belong. A person in the original position knows nothing about his physical or intellectual abilities, nor does he know his conception of the good or of his rational plan of life (137).

These persons in the original position "do not know, of course, what their religious or moral convictions are, or what is the particular content of their moral or religious obligations as they interpret them. Indeed, they do not know that they think of themselves as having such obligations" (206).

³ But this is not always the case (142).

Even though these people do not know what their beliefs are, or how those beliefs fare in society, or even that they hold any religious beliefs or obligations, they are not allowed to "take chances" with their beliefs by agreeing to the dominance of one religion in society, for "to gamble in this way would show that one did not take one's religion or moral convictions seriously" (207).

Rawls says candidly that "we want to define the original position so that we get the desired solution" (141).⁴ The "desired solution" of which Rawls is speaking is the acceptance of the two principles of justice that come out of the original position. These principles will now be stated briefly.

The original position is set up so that no one individual can win special privileges for himself and also so that he has no reason to accept any special disadvantages either. "Since it is not reasonable for him to expect more than an equal share in the division of social goods, and since it is not rational for him to agree to less, the sensible thing for him to do is to acknowledge as the first principle of justice one requiring an equal distribution" (150).

Thus Rawls' first principle of justice can be formulated as follows: 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 (302).

While the first principle of justice points toward equality, these may be cases in which inequality leads to the betterment of society in general and of the disadvantaged in particular (151). Considerations such as these lead to Rawls' second principle:

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 (302).

⁴ The original position developed by Rawls "has been defined so that it is a situation in which the maximin rule applies" (155). The "maximin rule for choice under uncertainty" applies to situations in which no rational estimate of probable outcome can be made, the person making the choice is concerned with a minimum standard of good, and certain alternative (152-54) choices have completely unacceptable outcomes.

III

I wish to consider, not the principles of justice he argues would be chosen in the original positions but rather the use that Rawls makes of the original position itself. It seems to me that Rawls places great importance on the original position in both determining (122-26) and justifying (17-21) the principles selected. To the extent that there are problems with the method of selection and justification for the principles of justice, the validity of those principles as a description of our sense of justice becomes problematic.

The two principles are justified because they would be the principles chosen in the original position; that is, these principles would be chosen as the principles of justice by hypothetical persons in a hypothetical social contract agreement (11-12, 118-21). These are the correct principles of justice because "their acknowledgement is the only choice consistent with the full description of the original position" (121). Rawls is firm and precise on this point: "If a conception of justice would be agreed to in the original position, its principles are the right ones to apply" (167).⁵

By the same token, other conceptions of justice, and their principles, must be rejected if they would not be accepted in the original position. For example, the principle of perfection, in any of its many forms (325), is rejected because it would not be accepted by the persons in the original position since it would violate the principle of equal liberty (327-28). As the interpretation of the original position changes, however, so do the principles of justice:

In order to arrive at the ethic of perfectionism, we should have to attribute to the parties a prior acceptance of some natural duty, say the duty to develop human persons of a certain style and aesthetic grace, and to advance the pursuit of knowledge and the

⁵ Rawls' use of the original position sounds like Kant's discussion of duty. "It is precisely in this that the worth of character begins to show—a moral worth and beyond all comparison the highest—namely, that he does good, not from inclination, but from duty" (Immanuel Kant, *Groundwork of the Metaphysic of Morals* [New York, 1964], p. 66). Rawls indicates that his principles of justice are categorical imperatives in Kant's sense (253). But the categorical imperative for Kant posits a universal law, universal for not only man but every rational being (Kant, 76, 98). Rawls' principles, however, are "universal" only in a universe based on his original position, and he himself admits that his principles would probably not be chosen from a complete list of alternatives (581). It is perhaps possible to read Rawls' use of the original position along the lines indicated by Kant: the principles are for "grounding philosophy on metaphysics," and the use of the original position is for "winning acceptance for it by giving it a popular character after it has been established" (Kant, 77).

cultivation of the arts. But this assumption would drastically alter the interpretation of the original position (328; cf. 121-22).

The principles of justice are justified through their attachment to the original position. Rawls' theory of justice then, relies in crucial aspects on the proper interpretation of the original position. Because of the importance of the original position, in developing this original position and the principles of justice which spring from it, "we should strive for a kind of moral geometry with all the rigor which this name connotes" (121).

The first problem that one faces in this task of moral geometry is that he is confronted with a jumbled maze of potential original positions: "There are indefinitely many variations of the initial situation and therefore no doubt indefinitely many theorems of moral geometry" (126).

How does one move from this mass of conflicting "initial situations" to the two principles identified above? According to Rawls, "for each . . . conception of justice there exists an interpretation of the initial situation in which its principles are the preferred solution" (121). Notice however that the two two principles of justice are determined, not by an "initial situation," but by the original position." One of many "initial situations" must somehow be converted into the "original position."

To arrive at the principles of justice one must select, out of all the possible alternatives, that description of the initial situation which will lead to them. This process is crucial, for a wrong selection of original position will lead to an incorrect set of principles, which will lead to an erroneous description of our sense of justice, which will lead in turn to an incorrect theory of justice.

There are two stages to the selection of and justification for choosing one of the initial situations as the original position: "The question of justification is settled, as far as it can be, by showing that there is one interpretation of the initial situation that best expresses the conditions that are widely thought reasonable to impose on the choice of principles yet which, at the same time, leads to a conception that characterizes our considered judgments in reflective equilibrium. This most favored, or standard, interpretation I shall refer to as the original position" (121).

Rawls speaks, first of all, of that interpretation of the initial situation "which best expresses those conditions that are widely thought reasonable." Later in the same paragraph Rawls calls this interpretation "phil-

osophically most favored" (122). This interpretation is, for Rawls, "the original position": "The concept of the original position, as I shall refer to it, is that of the most philosophically favored interpretation of this initial choice situation for the purposes of a theory of justice" (18).

The "philosophically favored interpretation" of the initial situation is the one which sets conditions that "are widely thought reasonable" (121), and in fact "must be characterized by stipulations that are widely accepted" (14; cf. 18, 584). Rawls does make an effort to show that his original position is made up of such widely accepted stipulations and states that "one should not be misled . . . by the somewhat unusual conditions which characterize the original position" (18).

The "somewhat unusual conditions" become obvious as Rawls develops this "widely accepted" original position. He states that it "seems reasonable and generally acceptable that no one should be advantaged or disadvantaged by natural fortune or social circumstances in the choice of principles" (18). He states that "it also seems widely agreed that it should be impossible to tailor principles to the circumstances of one's own case" (18). Notice the next sentence, however: "We should insure further that particular inclinations and aspirations, and persons' conceptions of their good do not affect the principles adopted" (18). Rawls does not claim that this idea is widely accepted, nor even that it is reasonable. "The aim," he says, "is to rule out those principles that it would be rational to propose for acceptance, however little chance of success, only if one knew certain things that are irrelevant from the standpoint of justice" (18).

Rawls sets this limitation, not because of its obvious reasonableness or its general acceptance. He sets this stipulation to insure that a "rational" alternative—an alternative which is potentially incompatible with his principles of justice—may not even be considered. Rawls states very clearly this method of operation: "We want to define the original position so that we get the desired solution" (141).

We now turn briefly to the most unusual of Rawls' "somewhat unusual conditions"; the veil of ignorance. "Among the essential features of this situation," writes Rawls,

is that no one knows his place in society, his class position or social status, nor does any one know his fortune in the distribution of natural assets and abilities, his intelligence, strength, and the like. I shall even assume that the parties do not know their conceptions of the good or their special psychological propensities. The principles of justice are chosen behind a veil of ignorance (12; cf. 136ff).

Rawls does not claim that the veil of ignorance is "widely thought reasonable," but that "the veil of ignorance is so natural a condition that something like it must have occurred to many" (137). Whether natural or unnatural, reasonable or unreasonable, widely accepted or widely rejected, Rawls' reason for positing the veil of ignorance is clear: "the arbitrariness of the world must be corrected for by adjusting the circumstances of the initial contractual situation" (141).

The argument for the original position based on wide acceptance, in a way begs the entire question. If one could arrive at the state of reflective equilibrium, and the principles of justice, without developing the original position, there would be no reason to introduce the notion of "original position." If, on the other hand, one needs this concept to assist one in constructing a theory of justice, wide or narrow acceptance seems to be immaterial. It is certainly immaterial if the acceptance required is that of men in general, for men in general do not have considered judgments on these questions (47-8). Perhaps however Rawls, by calling the original position the "most philosophically favored," intends to indicate that wide acceptance on the part of philosophers—men who do have considered judgments—determines which of all possible alternatives leads to the true principles of justice. This is, to say the least, a difficult position to hold, since the various alternatives under consideration stem from various of these same philosophers (123-24).

Rawls makes no effort to establish that his original position is "philosophically most favored" (nor does he explain exactly what "philosophically most favored" means). He posits this statement rather than proving it. He does not show that it is widely accepted as reasonable, but merely states that it is, which is not the same thing at all.

It is clear that Rawls produces no evidence to prove that there is general acceptance of the original position, which he has outlined as the reasonable position. To say this, however, is not to say that the original position is unreasonable. This question can be considered in two ways: (1) does the use of the original position argument add to the reasonableness of Rawls' case? ⁶ and (2) are the arguments adduced by Rawls reasonable, whether widely accepted or not? Simply put, given the framework which Rawls chooses, does he make a convincing case?

The question last raised is best considered by looking toward the second of Rawls' two methods of justification. We can justify selecting one interpretation of the initial situation as the original position by

⁶ I shall not consider this aspect of the question. See Ronald Dworkin, "The Original Position," *The University of Chicago Law Review*, XL (1973), 500-33.

choosing the one which "leads to our considered judgments in reflective equilibrium" (121).

The game works as follows: one determines which principles of justice one accepts ("considered judgments"), and then one builds an original position that will insure the desired outcome—"one tries to set out the different conditions embodied in the contractual situation in which their principles would be chosen" (122). One then justifies—"as far as it can be done" (121)—the selection of that particular initial situation because it coincides with his "considered judgments." Then one claims that whatever principles selected must be followed, for "if a conception of justice would be agreed to in the original position, its principles are the right ones to apply" (167).

Earlier I quoted Rawls as speaking about certain considerations "irrelevant from the standpoint of justice" (18). But to know what is "irrelevant" from that standpoint is to know something, perhaps a great deal, about justice. Rawls is involved in tailoring his original position to arrive at the principles he already accepts (141). If it is a problem of saving, simply "alter the motivational assumption" (128, 140). How can the principle of "equal liberty of conscience" be insured? Posit that one must take seriously their religious and moral beliefs, even if they do not know what those beliefs are (207).

Rawls begins his discussion of equal liberty of conscience by stating his conclusion: "The question of equal liberty of conscience is settled. It is one of the fixed points of our considered judgments of justice" (206). To requote a passage referred to earlier in this paper: "We want to define the original position so that we get the desired solution" (141).

If the claim of circularity is correct,⁷ then Rawls' argument from the original position ultimately fails. He recognizes this problem to some extent when he admits that his work is "highly intuitive" (121; cf. 124, 584). Perhaps in the long run all reasoning is circular,⁸ with the difference being the diameter of the various orbits. More important that the charge of circularity is the charge of sterility, which will now be briefly examined.

⁷ Other problems of circularity arise in Rawls' theory. In his theory, he says, "the concept of right is prior to that of the good" (31). Yet later Rawls writes that "something is good only if it fits into ways of life consistent with the principles of right already on hand. But to establish these principles it is necessary to rely on some notion of goodness, for we need assumption about the parties' motives in the original position. Since these assumptions must not jeopardize the prior place of the concept of right, the theory of the good used in arguing for the principles of justice is restricted to the bare essentials" (396). Rawls call this the "thin theory" of justice. Thick or thin, it still presents a problem for Rawls. Also see 433-34.

⁸ *Republic*, 533 b-d.

IV

What began as an effort to describe "our sense of justice" has arrived at the point of "a purely hypothetical situation" which will be used to explain our very real sense. Rawls indicates the reason for this in a comparison of moral theory (of which the theory of justice is a part) with social theory:

For while the theory of price, say, tries to account for the movements of the market by assumptions about the actual tendencies at work, the philosophically favored interpretation of the initial situation incorporates conditions which it is thought reasonable to impose on the choice of principles. By contrast with social theory, the aim of moral theory is to characterize this situation so that the principles that would be chosen, whatever they turn out to be, are acceptable from a moral point of view (120).

Moral theory is not concerned with "the actual tendencies at work" in the normal, everyday, judgments discussed earlier. Moral theory is concerned simply with what Rawls calls "a moral point of view." It is such a theoretical position that allows Rawls to "look for a conception of justice that nullifies the accidents of natural endowments and the contingencies of social circumstance" because they are "arbitrary" (15).

While there are perhaps many difficulties with this position, I want to indicate now only one. Principles of justice that are derived from abstract reasoning that is not tied into actual human life, activity and thought may in some way become highly questionable when applied to that life, activity, and thought.

Rawls holds that actual institutions are to be judged in light of the principles of justice derived from the original position, and that those actual institutions are to be considered unjust insofar as they depart from the ideal conception of justice (246). But Rawls is forced to admit that in fact this position is highly problematic. Speaking of his ideal conception of justice, Rawls states:

By putting these principles in lexical order, the parties are choosing a conception of justice suitable for favorable conditions and assuming that a just society can in due course be achieved. Arranged in this order, the principles define then a perfectly just scheme; they belong to ideal theory and set up an aim to guide the course of social reform. But even granting the soundness of these principles for this purpose, we must still ask how well they apply to institutions under less than favorable conditions, and whether they provide any

guidance for instances of injustice. The principles and their lexical order were not acknowledged with these situations in mind and so it is possible that they no longer hold (245).

To state that Rawls does not follow Plato is not to refute Rawls' position. But it is to indicate that his work does not have the political vitality that is found in the *Republic*: the understanding of how political men and institutions act. While political philosophy cannot end with the city, that is where it should at least begin. Rawls, to an extent, recognizes this: "The fundamental principles of justice quite properly depend upon the natural facts about men in society" (159). He begins his book by reflecting that "a society is well-ordered . . . when it is effectively regulated by a public conception of justice" which "everyone accepts" (4-5). He also indicates that "existing societies are . . . seldom well-ordered in this sense, for what is just and unjust is usually in dispute" (5).

In addition, Rawls would like for his work to move the existing political structure toward a more just structure (246, 303). But, by identifying important political considerations as irrelevant and arbitrary aspects of the social world from a moral point of view (18), Rawls effectively cuts off his moral theory from the real political groundings that it needs. Rawls argues that "the arbitrariness of the world must be corrected for by adjusting the circumstances of the initial contractual situation" (141). But, in arbitrarily adjusting his theory of justice to fit his considered judgments, he removes himself from the political *Lebenswelt*, which is irrevocably tied to the "arbitrariness of the world."

But the relation of the political philosopher to politics has long been problematic.⁹ The political philosopher has always had an easier time communicating with other minds than with political men.¹⁰

Rawls attempts to justify his original position as the "one interpretation of the initial situation which best expresses the conditions that are widely thought reasonable" (121). This argument is a rhetorical effort to communicate with political men. The serious argument made by Rawls, as he himself acknowledges, is that his interpretation is reasonable (582), and this argument communicates with other minds. Unfortunately Rawls stipulates rather than proves this reasonableness.

Rawls states early on that, "for the purposes of this book, the views of the reader and the author are the only ones that count" (50). Just

⁹ *Republic* 488a-489d, 519b-521d.

¹⁰ Leo Strauss, *Persecution and the Art of Writing* (Glencoe, Ill.: The Free Press, 1952), pp. 22-37.

as moral philosophy is Socratic (48), so is Rawls' book an attempt at Socratic education: the opening of a dialogue between two minds. If, in the long run, Rawls fails in this effort, it is because that, while he remembers the Socratic principle that "the fundamental principles of justice quite properly depend upon the natural facts about men in society" (159), he does not remember it enough.