

Rawls and Rousseau on the Social Contract

JOSEPH M. GRCIC
Loyola University of Chicago

In his A Theory of Justice, John Rawls claims his social contract theory can be considered part of the social contract tradition which includes Locke, Rousseau and Kant.¹ The purpose of this essay is to determine what defines a tradition such as that of the social contract and to evaluate Rawls' claim that he is in the same tradition as Rousseau. At the same time, I analyze the general nature of the contract model and so clarify some of the controversial points in Rawls' and Rousseau's political theories.

I. Rousseau:

In the Social Contract Rousseau characterizes the contract in these terms: "although they (the clauses of the contract) have perhaps never been formally set forth, they are everywhere the same and everywhere tacitly admitted" (14).² Rousseau believes these clauses can be reduced to one: "the total alienation of each associate, together with all his rights, to the whole community" (14). Each must alienate all rights since if any right were left to individuals, without a "common superior" to decide between conflicts of individuals and public, the state of nature would ensue and the association dissolve. Each participant,

in giving himself to all, gives himself to nobody; and as there is no associate over which he does not acquire the same right as he yields others over himself, he gains an equivalent for everything he loses, and an increase of force for the preservation of what he has. (14)

In another definition of the social contract, Rousseau introduces the key concept of the "general will" (*volonté generale*) (henceforth abbreviated GW):

Each of us puts his person and all his power in common under the supreme direction of the general will and in our corporate capacity, we receive each member as an indivisible part of the whole. (15)

In the Emile, Rousseau is even more insistent on the radical nature of the alienation demanded by the contract:

As an individual everyone of us contributes his goods, his person, his life, to the common stock under the supreme direction of the general will.'

The body politic of the sovereign, then, derives its existence only from the contract. The sovereign, being only the aggregate of the individuals, is the supreme power in the society.

Few concepts in the history of political theory have been as controversial as Rousseau's notion of the GW. There is no question, however, that it is the key to understanding Rousseau. According to Rousseau, each individual has a particular will which aims at his particular interests. In the state of nature only the particular will and private interest exist. The GW, which exists only in the political state, aims at the common or general interest. For Rousseau, then, the 'will' is distinguished by its object; the private will tending towards the private interest, the general will at the general object of interest. Sovereignty is nothing else but "the exercise of the GW" (23). Moreover, Rousseau contends that the GW "is always right" (26), but he does not take this to mean that the majority is always right. Rousseau wants to draw a distinction between the GW and the "will of all," the latter being simply the sum of particular wills (26). How then is the GW ascertained if not by a majority vote?

Rousseau's answer is as complex as it is elusive. At first he suggests that the GW remains as "sum of the differences" between the particular wills (26). More specifically, the GW emerges from the deliberation of all the citizens when they have the "adequate information" and if the "citizens had no communication with one another" (27), i.e., when there are no factions or parties within the state which seeks their own ends rather than the common end. Though Rousseau rejects the view that the GW is simply the aggregate of particular wills, he does believe it can be expressed by a majority will under certain conditions. The majority will is the GW if when the issue is before the national assembly

the people is asked not exactly whether it approves or rejects the proposal, but whether it is in conformity with the GW, which is their will . . . and the GW is found by counting votes.
(106)

If one finds oneself in the minority, then Rousseau believes one must conclude "I was mistaken" about the GW and one's own will. In the above quotation Rousseau

states the GW "is their will," i.e., the will of the individual; does then each person have two wills, the general and the particular? No, what Rousseau must mean is that the GW is the will of the individual as citizen, not as a private individual. Hence if one is in the minority on a given vote, one is in error as to one's interest as citizen, though not necessarily as private person.

Once the contract is completed, the individual parties to it have renounced all rights and possessions to the body politic. For Rousseau, there are in principle no limits to the GW; "the social compact gives the body politic absolute power over all its members" (28). If an individual finds oneself in disagreement with the GW, he will be "compelled to do so by the whole body. This means nothing less than that he will be forced to be free" (18). If 'freedom' is construed as the absence of external coercion, then "forcing" someone to be free is a contradiction in terms. But another interpretation of this Rousseauian definition of freedom can make this notorious passage consistent. In the state of nature man enjoyed what Rousseau calls "natural liberty" (19), i.e., the unlimited right to do anything and have anything instinct suggested without coercion from others. In civil society one loses natural liberty but gains civil or "moral liberty" (19). This is freedom to do what is morally permissible as defined by the GW. Moral liberty "alone makes (man) truly master of himself" (19). Doing what one wants, or natural liberty, is to follow "the mere impulse of appetite (which) is slavery" (19). Moral liberty is desiring to do what one ought to do, not merely want to do, which only "obedience to the law" can reveal as the expression of the GW. Hence Rousseau saw freedom as doing what one ought to do, or being unrestrained from doing what one ought to do; if one erred as to what one's true moral duty was, one could be forced to be free, i.e., forced to do what is really in one's interest, as opposed to doing what is only apparently one's interest.

II. Rawls:

Rawls believes his principles of justice would be chosen by free, rational, equal and self-interested individuals within a well-defined context which he calls the "Original Position" (henceforth, OP). As in Rousseau, this agreement is not an historical event but a hypothetical construct; it consists in the claim that 'individuals' or persons of the OP (henceforth, POPs) would choose certain principles of justice under properly defined circumstances. These circumstances which define the OP consist of what Rawls takes to be certain rules of rationality, general empirical facts and relevant moral intuitions.⁴

The "Veil of Ignorance" expresses some of the conditions that members of the OP are subject to. This fictional veil means POPs are ignorant of: a) their place in society, status or position; b) fortune, natural talents such as intellectual ability or physical strength; c) their particular conception of the good; d) their psychological propensities; e) to which generation they belong. These conditions allow for the principles of justice to be chosen without the influence of prejudice that knowledge of one's natural and social circumstances would lead to. To allow information about such matters of social happenstance would be to select principles which would be shaped, at least in part, by this 'prior' society and its social contingencies which may be an unjust society.

This way of regarding justice Rawls calls "justice as fairness" for these principles would result from free agreement among individuals under the above described circumstances. For Rawls, a practice is fair when none of those participating in it feel he is being taken advantage of or being compelled to give in to what he considers illegitimate claims.

Having described the limits on the knowledge of these hypothetical persons, Rawls must next give these individuals some content to explain their decision in favor of some principles of justice rather than others. By stipulating that these decision-makers are "free" Rawls means they have no authority over one another. By "rational" he means they are interested in furthering and taking the most effective means to their given ends (14). By "equal" is meant POPs are capable of having a conception of the good and a sense of justice and, in addition, they have the same degree of decision-making powers in the OP (19). Finally, by adding that POPs are "self-interested" Rawls does not intend to convey that POPs are egoists, but merely that they are concerned in advancing their goals and desires.

Given these conditions, Rawls believes the principles of justice chosen would be:

- I) 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.
- II) 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)

By "basic liberties" is meant the right to vote, hold public office, free speech, assembly, thought and property (61). Rawls believes these two principles are a special case of a more "general conception" of justice, namely:

. . . all social values, 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 these values is to everyone's advantage. (62)

The reason POPs would choose these principles is that they provide for what Rawls calls "primary goods" or "what a rational man wants whatever else he wants" (92). That is, they are necessary means to fulfilling one's plan or goal of life, whatever these happen to be. These goods include rights and liberties, opportunities and powers, income and wealth and self-respect. The two principles would be chosen by rational self-interested persons as the best way for each to secure his ends. Though POPs do not know their particular conception of the good they do know they have a rational life-plan and to achieve this end they prefer more primary social goods which the two principles of justice guarantee.

III. The Logic of the Contract:

The theories of Rousseau and Rawls, and that of Locke and Hobbes as well, are complex structures made up of what may be termed formal and material components. Contract theory is distinguished from other models of political obligation in that agreement or consent is the foundation of any legitimate political authority; this is the formal dimension which all contract theories share. In Rawls, the agreement is hypothetical, i.e., what properly defined individuals would agree to in the OP. The hypothetical contract obligates, Rawls believes, just as a real contract does because it is a result of moral beliefs we in fact possess or can be persuaded to hold after proper consideration. Rousseau's contract also has the non-historical or hypothetical form for, as has been pointed out above, Rousseau concedes the contract as he describes it may "have perhaps never been formally set forth" (14). (Indeed, considerations to follow will show that the contract cannot be other than a hypothetical one.) However, this similarity of the formal component is only partial, for the hypothetical aspect of Rousseau's contract has an actual element which is essential for its completeness.

Recall, in the Social Contract, Rousseau first stipulates that for the GW to emerge from an assembly of citizens three conditions must be met: 1) relevant

information must be available (27); 2) citizens must not form parties or communicate with one another (27); and 3) the question must be put to the assembly in this form: "Is it in the common interest?" and the decision makers must ignore their own private interests. Now, this closely approximates Rawls' Original Position; the last two conditions correspond to the Veil of Ignorance and the formal constraints of universality and generality on principles of justice and the first to the general knowledge people in the Original Position are said to have at their disposal. The crucial difference, however, is that Rawls' OP is a conceptual device constructed so as to lead to the desired conclusion, the Two Principles, whereas Rousseau's recommendations are for a real assembly of people who know who and what they are and what they want. This difference is not without significance.

The problematic character of the idea of the GW is revealed when the presuppositions of the three conditions that the General Assembly must meet to derive the GW are made explicit. First, to disregard one's own particular interests assumes a kind of altruistic moral perfection, which Rousseau cannot explain or guarantee. But even if they were to disregard their interests, this would presuppose that some sort of cognitive perfection has been attained, for they must know infallibly what their interests are if they are to effectively disregard them. Even granted that such a state were attainable, it would still be a private psychological state for which there seem to be no public criteria and therefore no way to determine whether the condition is in fact present. The second condition states that the Assembly must have all the "relevant information" (27) but Rousseau gives no criteria for this relevance. The third and final pre-condition for the Assembly requires that there be no factions or parties formed. Clearly, this condition, too, is unattainable for it would necessitate an omnipresent and omniscient state to monitor all communications between individuals at all times. In sum, to assert, as Rousseau does, that the GW is infallible (26), is to make the implausible claim that human society composed of fallible and finite creatures can reach political infallibility in history.

The above difficulties with the GW are compounded by the fact that the GW has only a procedural interpretation, i.e., it is the outcome of a procedure followed by the General Assembly; there are no independent substantive criteria for judging whether the Assembly's decisions are in fact the GW. Using Rawls' terminology, Rousseau's method is an instance of "pure procedural justice" (86), which obtains when there is an appropriate procedure which, if followed properly, will yield the correct or just result, but there is no independent criteria with which to evaluate the outcome. As in games of chance, so too in the case of the GW,

the procedure must actually be carried out before the solution can be known. In Rawls, however, there is a relatively independent set of criteria by which to adjudicate the outcome of his procedure, our considered moral judgments. Thus, Rawls' approach may be termed what Rawls calls "imperfect procedural justice" because criteria of the correct outcome do exist. In imposing conditions that can only be applied in a hypothetical situation on an actual collection of persons, Rousseau has outlined a decision procedure which cannot be actualized in the real world. But without the General Assembly to complete and define the nature of the political community which the General Will demands, Rousseau's contract remains an empty abstraction.

This limited concurrence in the definition of the formal component of the contract in Rawls and Rousseau is further diminished when the variables that constitute the material components of contract theory are defined. Agreement never occurs in a vacuum nor is the contract some eternal unchanging form; agreement always results in a specific contract whose specificity is determined by who agrees to what and for what purpose. The manner in which these material components are characterized will reveal the sources of the incommensurability between Rousseau and Rawls.

What contractees agree to, the clauses that constitute the contract, is a function of some goal or purpose they wish to actualize. These goals may be values, institutions, relationships, conditions or, more generally, a certain kind of socio-political structure or community. What these goals are is a function of three factors: 1) what the contractees perceive as actual, i.e., their circumstances and condition; 2) what they believe is possible for them to achieve given man and his condition; 3) what the contractees see as desirable, the values or rights they wish to promote or preserve.

In Rousseau, the purpose of the contract was not the protection of rights, but the more nebulous one of establishing a society where the values of liberty and equality would be realized and promoted through the General Will and the total alienation of rights. But, more importantly, Rousseau's ultimate goal, as will be shown below, was the creation of a new kind of human nature, a nature more noble, compassionate and wise.

In the case of Rawls, the goal of his theory, from our own point of view, is to clarify our moral intuitions, but from the point of view of the POPs, it is to define principles of justice to ensure themselves the basic goods, namely, rights, liberties, opportunities, power, income, wealth, and self-respect. And to achieve this, they agree to the Two Principles.

The material category of Agreement 'by whom' is concerned with the definition, essence or nature of the contractees, those party to the contract. This is a

crucial component of any political theory, including that of the contract, for the contract consists in the definition of the relations that must exist between those party to it and the nature of the relations must be at least partly based on the nature of the relata, the contractees. In establishing the nature of man three components are especially relevant here: 1) the character and extent of his innate capacities, motivations and weaknesses; 2) the relation and influence of the physical and social environment on his condition and achievements; 3) the nature of his relation to other persons. These elements are relevant in determining the degree and nature of the obligations, functions and responsibilities of the contractees.

Though we cannot here embark on a comprehensive and fully detailed discussion of Rousseau's concept of man, still a brief discussion is inevitable for an understanding of his contract.

It is clear from the First Discourse (FD) and Second Discourse (SD) that Rousseau believes contemporary man has degenerated and become decadent, immoral, in short, alienated from his essential nature.⁵ Rousseau's theory of ideal human nature has at least five essential elements. First, Rousseau tells us that the two basic principles of man's essence are those which concern our well-being, namely, "self-preservation" (SD, 95) or "Amour de soi," and the principle which "inspires in us a natural repugnance to see any sensitive being perish or suffer, principally our fellowmen" (SD, 95, 130-1), namely sympathy, pity or compassion. The third essential component of man is his freedom, as Rousseau says in his famous lines of the Social Contract, "Man is born free and is everywhere in chains" (3, 19). And in the SD, it is man's freedom more than his "understanding" or reason which constitutes man's distinction from other animals (114). This freedom means man, unlike other animals is not a slave to his instincts but "is free to acquiesce or resist" (SD, 114) his instincts. In addition to this ontological sense of freedom, according to Rousseau, man in the state of nature has what may be called "economic" and "psychological" freedom. By the former, Rousseau means for a man to be fully free, and hence fully human, he must have enough to satisfy his needs so that he does not become a slave to the rich.⁶ Psychological freedom expresses Rousseau's belief in the need for the absence of dependence on others for one's sense of worth or respect; recall how in the SD Rousseau laments man's alienated state where he comes to "live in the opinion of others."

The fourth dimension of man's nature is his ability and need to labor to satisfy his wants. Rousseau does not state explicitly that this is part of man's essence, but it is implicit in his writings. We saw that one of his basic critiques of the then contemporary

French society in the First Discourse is its luxury and idleness (46, 49, 53). In the Discourse on Political Economy Rousseau contends society must avoid extreme wealth so that "Labor is always necessary and never useless for its acquisition" (255). And, finally, Rousseau's objections to commerce and his advocacy of agriculture and the barter system, is based on his understanding that entrepreneurs are idle and only farmers are said to really earn their income.

The fifth element of man's essence Rousseau calls man's "perfectibility" (SD, 115), or the faculty which "successfully develops all the others" (SD, 114). Unfortunately, Rousseau says little beyond this to clarify the nature of this supposed faculty. In addition, he rejects "reason" as essential to man, which seems the only faculty that "perfectibility" may name. Such remarks as "the state of reflection is a state contrary to nature and the man who meditates is a depraved animal" (SD, 110) and "reason engenders vanity and reflection fortifies it" (SD, 132) supports the reading that reason is not part of perfectibility. However, in the First Discourse, which is often interpreted as the apotheosis of ignorance and barbarity, Rousseau does in fact allow for the exercise of the intellect and the development of science and the arts as long as their influence consists in increasing the "happiness of the people" (63). This can occur only when governmental power and man's wisdom work together, not separately and at odds with one another. Hence, it seems reason, correctly used, may indeed have a proper role in the ideal society of men.

In sum, man properly understood, is a creature which preserves itself, has compassion for others, is free, labors and is perfectible (rational). Rousseau objects to man's present condition because it frustrates his nature in that many individuals are rich and idle, and many unfree due to their poverty, and perhaps most significantly, most have lost the virtue of compassion which has hurled society into a veritable state of war. In a corrupt and alienated society "amour de soi" or self-preservation degenerates into "amour-propre" or conceit and vanity. Rousseau believes his social contract reestablishes freedom and compassion by totally alienating all rights, rediscovers the virtue of labor by recreating an agrarian city-state, and baptizes the arts of sciences which now work for man, not against him.⁷

Rawls' own views on the nature of man are not easy to ascertain for he nowhere explicitly states them. Yet his descriptions of the POPs do give some indication of how he sees man. He sees him as morally and intellectually limited, free, self-interested, i.e., concerned to promote his own good, rational and non-envious, since envy tends to make everyone worse off and is therefore irrational. Of all these traits,

self-interestedness is the most crucial component of Rawlsian man. It is on that basis that POPs choose to maximize the "primary goods" or "what a rational man wants, whatever else he wants" (92). On this view, rationality logically entails self-interestedness, and without it, the principles of justice POPs would choose would be substantially different. This self-centered view of man is further revealed by man's existence under what Rawls calls "circumstances of justice" (126) including scarcity of basic goods, which, in turn, adds to the potential for conflict.

Of course, Rawls is quick to note that the motivation of POP is not necessarily that of real people outside the OP (147-8), yet if there is no structural isomorphism between the relevant motives within and without OP, then the two principles of justice would be irrelevant to man as he is. In this regard, it is important to mention the Aristotelian Principle, which Rawls doesn't discuss until the end of his treatise, is an important element of human nature which apparently, POPs do not know about and hence is not a contributing factor in their deliberation.

The most fundamental reason for the difference between Rousseau's and Rawls' political theories lies in their definition of what man is. As we noted above, Rawls assumes that human nature as it is, with the tendency to self-interest, was acceptable and a rational starting point from which to build a theory. Rousseau, on the other hand, sought to destroy the roots of self-interest in man for he saw it as a source of decadence, disunity and ultimately war itself. This is quite apparent especially in Rousseau's Second Discourse. There Rousseau observes that man's essential trait of sympathy and compassion was overcome and suppressed by the development of private property and the ensuing competition for superiority which brought forth envy and vanity among men. As Rousseau saw it "in a word, competition and rivalry on one hand, opposition of interest on the other and always the hidden desire for profit at the expense of others. All these evils are the first effect of property and the inseparable consequence of nascent inequality." This eventually resulted in a state of war.

We see then, that the similarity between Rawls' and Rousseau's respective theories of the social contract rests simply on the formal dimension of agreement as the necessary condition for the establishment of political authority; they diverge on who (human nature) makes the agreement, and for what purpose (their visions of the "good" society). The society which Rousseau envisioned resulting from and constituted by the social contract was radically distinct from the society in which he saw contemporary man. It would be a society populated by a new type of man, a type which had overcome vanity, competition, pride and self-

interest. At the heart of Rousseau's theory, then, is a revolutionary interest to transform mankind and society. This approach is diametrically opposed to that of Rawls.

Rawls assumes that our present considered moral judgments and given human nature are an acceptable starting point for political theory. To be sure, the OP does make more precise these moral intuitions, but the changes required by the Two Principles can be carried out in an evolutionary manner within the existing institutions of western liberal democracy. This is made even more perspicuous by Rawls in a recent article where he states:

. . . we are not trying to find a conception of justice suitable for all societies regardless of their particular social or historical circumstances. We want to settle a fundamental disagreement over the first form of basic institutions within a democratic society under modern conditions . . . Our hope is that there is a common desire for agreement as well as a sufficient sharing of certain underlying notions and implicitly held principles . . . the aim of political philosophy . . . is to articulate and to make explicit those shared notions and principles . . .'

Ultimately, then, what lies at the heart of the disagreement between Rousseau and Rawls is radically opposed visions of what man and human society should be; unfortunately, it is not clear that these respective visions are ultimately persuasive.

NOTES

¹John Rawls, A Theory of Justice, Cambridge: Harvard University Press, 1971.

²All quotations from The Social Contract are from G. D. H. Cole's translation, New York: Dutton, 1950. Numbers in parentheses refer to pages of this work unless otherwise indicated.

³J.J. Rousseau, Emile, B. Foxley, trans., New York: Dutton, 1938, p. 424.

⁴In this section, numbers in parentheses refer to John Rawls' A Theory of Justice, op. cit.

⁵J.J. Rousseau, The First and Second Discourses, R.D. Masters, ed., New York: St. Martin's Press, 1964.

⁶J.J. Rousseau, Discourse on Political Economy, in The Social Contract, New York: Carlton House, p. 251.

⁷Cf. J.J. Rousseau, Considerations on the Government of Poland, pp. 226-28, and Project for the Constitution of Corsica, p. 291, in Rousseau: Political Writings, E. Watkins, ed., New York: Nelson, 1953.

⁸R.D. Masters, op. cit., p. 156.

⁹John Rawls, "Kantian Constructivism in Moral Theory," Journal of Philosophy, September 1980, p. 518.