



## **Rawlsian Affirmative Action: Compensatory Justice as Seen from the Original Position \***

*Robert Allen*  
[RAllen8219@aol.com](mailto:RAllen8219@aol.com)

**ABSTRACT:** In *A Theory of Justice*, John Rawls presents a method of determining how a just society would allocate its "primary goods"-that is, those things any rational person would desire, such as opportunities, liberties, rights, wealth, and the bases of self-respect. (1) Rawls' method of adopting the "original position" is supposed to yield a "fair" way of distributing such goods. A just society would also have the need (unmet in the above work) to determine how the victims of injustice ought to be compensated, since history suggests that social contracts are likely to be violated. This paper is an attempt to determine the remedial measures that would be selected using Rawls' method. I contend that only two of the three most widely used "affirmative action" policies would be selected from the original position. I also sketch another compensatory policy that would pass Rawls' fairness test.

### I.

Affirmative action is public policy designed to compensate the victims of injustice. (2) To be thus disadvantaged, in Rawls' scheme of things, is to have suffered in some way from having had less than one's fair share of the primary goods (62). This measure, according to Rawls, ought to be determined by the two principles that would be selected in the original position (17-22). The "first principle," which is "lexically prior" to the second, dictates that each member of society be granted every shareable personal liberty, a liberty being shareable just in case one's exercising of it would not prevent others from doing so (60-1, 250). The "second principle" states that the other primary goods are to be distributed in an egalitarian fashion unless an unequal distribution would leave everyone with more of these goods than they would otherwise have. In addition, it mandates that every advantage is to be the reward of holding a position for which all members of society have the opportunity to compete (60-1, 84, 302-3). (In what follows, this caveat will prove critical.)

I shall construe the goal of affirmative action policies in Rawls' terms. Thus, they are to be understood as being designed to secure for those who were harmed by a violation of either one of the above principles the measure of primary goods to which those principles entitle them as well as a remedy for the effects of their deprivation. They, then, require us not only to change current distributional practices so as to make them fair, but also to help those disadvantaged by past injustices become capable of utilizing the assets such a change would bring their way.

One way of achieving this goal is to mandate that the administrators of educational, governmental, and business institutions employ aggressive means of recruiting the victims of injustice. This practice, hereafter "AA1," is supposed to insure that such persons, who might otherwise be ignorant of the academic and employment opportunities thereby offered, are made aware of their existence and encouraged to take advantage of them.

The second form of affirmative action dictates that a member of a disadvantaged group be given a sought after position in the event that she is as qualified for it as other aspirants. The idea here is to compensate a victim of injustice by having her disadvantaged status serve as a "tie-breaker." The justification typically offered for this policy, hereafter "AA2," is that any victim of injustice who 'holds her own' against those who possess their fair share of the primary goods would, *ceteris paribus*, be better qualified had her victimization not occurred. Such a person, it seems reasonable to suppose, would have outdone her rivals but for her disadvantage, this having kept her from more fully realizing her potential. (There are those who question this assumption, maintaining that in these matters one can only speculate as to how things would have played out counterfactually. Yet it seems hard to imagine what a disadvantage would be here if not something that *causes* an individual to do less well in comparison to her peers than she could have but for...?)

Finally, there is the version of affirmative action that allows the use of college admission and hiring "quotas." This practice, hereafter "AA3," would provide redress to the members of disadvantaged groups by setting aside for them a fixed number of "coveted" positions (these being the positions such as that of a university student or a partner in a prestigious law firm, to which the advantages allowed by the second principle are affixed). Included here are also policies, which in lieu of set asides (which are currently illegal) place the victims of injustice at a competitive advantage over other aspirants to these positions, allowing them to be filled by individuals whose qualifications are not higher than or equal to those of their rivals. AA2 is meritocratic AA3 is not.

Turning to Rawls' method, to adopt the original position is to place oneself behind a "veil of ignorance": it is to discount one's particular needs, interests, abilities, gender, race, religion, and socioeconomic status- in short, personal characteristics that are not shared by all others. One may regard oneself only as an arbitrary member of the society whose distributive principles are to be determined. Thus, in selecting distributive principles from the original position, one is not allowed to employ considerations the recognition of which would tend to make one partial towards those who share some or all of one's personal characteristics. The veil of ignorance prevents bias from influencing one's decision making (12). That is not to say that self-interest will not play a role. It will be limited, however, as follows.

Adopting the original position, Rawls argues, guarantees the selection of "fair" distributive principles (12-13, 17-19). It does so by compelling one to opt for principles whose application would provide *each* member of one's society with enough of the primary goods to realize her "plan or life." Being "ignorant" to which groups one belongs, one would not, as the rational agent Rawls presumes one to be, select principles whose application would leave some members of one's society disadvantaged. For to do so would be to put *oneself* at risk of being unable to realize one's goals, a risk to which a rational agent would be averse: for all one knows behind the veil, one may be a member of any disadvantaged group one could form. Self-interest would thus dictate that one not form such a group, mandating instead that one secure for each member of society enough of the primary goods to realize her plan of life. A just distributive principle, according to Rawls, is such that a rational agent would choose it if the operation of his self-interest were limited in this manner, i.e., if she were being fair.

As noted above, Rawls believes that in the original position a rational agent would choose the first and second principles of justice. He leaves unanswered, however, the question of how to make amends to the victims of injustice (145). Foreseeing the possibility of the above principles being breached, I propose to answer this question by choosing, from behind the veil of ignorance, principles of compensatory justice. I begin by determining whether or not AA1-3 would be acceptable from this perspective.

## II.

Would AA1 pass Rawls' fairness test? To answer this question, it must be determined whether or not AA1's application would have an unfair "downside." Staying within Rawls' system, I shall stipulate that the employment of a principle of compensatory justice would have an unfair downside (that is, expected consequences that would warrant the principle's rejection in the original position) just in case it would involve the violation of either one of the original distributive principles. In choosing principles of remedial justice, the first and second principles of justice are to act as constraints, besides those imposed by the veil of ignorance, on what measures may be chosen. (Though it turns out below that it is the second principle that is most germane to the matter being considered, the first principle and the veil of ignorance still play roles here, the former functioning to protect liberties the latter to insure that no one disadvantaged group is favored over another.) This proviso is in keeping with Rawls' assumption that those adopting the original position are committed to upholding the terms of whatever agreement is thereby reached (145).

There are parties to the current debate over affirmative action who might challenge this stipulation, maintaining that those who gained from discrimination *before* the principles were in place should expect to suffer a measure of injustice until their victims have been fully compensated. Whence comes the obligation to be fair to such persons (or their beneficiaries), especially if respecting their rights stands in the way of redressing those whose rights they have disregarded? But while fairness surely dictates that the beneficiaries of injustice abdicate their ill-gotten gains, it does not require additional sacrifices on their part. In particular, it does not require that they forfeit any *rights* others possess. To expect them to do so, is only to invite further claims of injustice, preventing the establishment *once and for all* of a 'level playing field'. The process of selecting principles of justice is rendered pointless unless it is assumed that its outcome will lead to as much.

To the objection that the issue of remedial justice belongs to "nonideal" theory (245-6) and is thus not to be settled by appeal to the first and second principles, I would respond that "strict compliance" or "ideal" theory does establish limits to what can be done by way of compensatory policies. It seems to me that the principles selected in the original position are not to be violated unless it can be shown that their suspension is entailed by the *only* effective means of redressing injustice, which, as will be demonstrated, is not the case. Taking them seriously would appear to require acknowledging that even those who may be the beneficiaries of injustice have rights that must be respected. Surely there is a high burden of proof that is to be met before it is warranted to fight injustice with injustice. Rawls' remark that the principles of justice are "generally relevant," (246) at least suggests such conservatism.

Our reluctance to breach the principles we had established would be bolstered by the realization that it may be very difficult to single out those whose talents, abilities, and qualifications are the by-products of injustice. While a case might be made for requiring extraordinary sacrifices of such persons, it is just not obvious in many instances where natural advantages end and those that stem from wealth and privilege begin. Can we say, e.g., that Wittgenstein would not have become a great philosopher but for his father's exploitation of his workers? Conversely, while it seems plausible to assume, as below, that rough equals would have been distinguishable had one of them not been harmed by

injustice, it is a stretch to suppose a lesser qualified candidate would have outdone her rivals but for her disadvantaged background. Without being able to draw sharp lines here, we run the risk of harming persons who have not profited from injustice and making amends to those who are not really its victims. Better then to be certain that all other compensatory policies are inadequate before nullifying our original contract. (3)

The question, then, we must ask of a proposed means of redressing injustice is: would it lead to some members of society being deprived of the measure of primary goods to which they are entitled according to the first and second principles of justice?

With respect to AA1, the answer to this question is no. Here one must be wary of violating the second principle's provision regarding "fair equality of opportunity" (84). But the aggressive recruiting mandated by AA1 would not entail that there be fewer coveted positions open to some persons than there are to others. It would simply bring into the field of competition individuals who would otherwise be left out. (I am assuming that the positions in question are existing ones and not ad hoc openings tailored exclusively to disadvantaged candidates. The creation of such positions falls under AA3.) Thus, assuming that AA1 would not breach other provisions of the two basic principles- that is, would not lead to inequalities in the distribution of personal liberties or the other primary goods (it being presumed that those who are not from disadvantaged groups are already aware of the opportunity to compete for existing coveted positions)- it would be acceptable to someone in the original position.

Jan Narveson raises the following objection to this conclusion. (4) It follows from a general concern over the measure of primary goods to which each member of a Rawlsian society is entitled. Narveson presents the defender of Rawls with a dilemma, the upshot of which is that we owe each other nothing but liberty, so that a victim of injustice is, at most, entitled to lost wages. Requiring employers to aggressively recruit the victims of injustice would unjustifiably infringe upon the formers' economic liberty.

The first horn is that, given the lexical ordering of the principles, there is no case in which applying the second principle is justified. (5) If we take seriously Rawls' stricture that liberty "can be restricted only for the sake of liberty," then it seems that assistance to the disadvantaged must not entail depriving someone of an opportunity of becoming wealthier, e.g., by disallowing, in the name of AA1, an employer's foregoing of a costly search for disadvantaged job candidates. The absolute priority of the first principle appears to make it impossible to promote equality by invoking the second.

Narveson realizes that a Rawlsian can evade this horn by pointing out that the first principle covers only the sort of liberties enumerated in the Bill of Rights. (6) It is not intended to apply to one's business dealings. But even if such activities did fall under the first principle's scope, the practices that Narveson mentions would not thereby be protected. The liberties they entail are not shareable: allowing employers to hire whomever they please licenses them to confine the members of a racial minority, say, to low-paying jobs, thus preventing them from becoming employers themselves. The first principle would place restrictions upon hiring practices.

The second horn, according to Narveson, is inescapable. There is no way to *justify* inequalities as beneficial to the least advantaged, as Rawls attempts to do with the difference principle. That many persons will not perform a demanding job without being given a higher salary than others, does not warrant having inequalities. It only "excuses" them. The difference principle cannot be used to license economic incentives involving inequalities. Rawls is left with the strict egalitarianism he sought to avoid. (7)

The Rawlsian may respond, however, that it is not the individual's *taking* of a higher salary that the difference principle is supposed to justify. Rather, it is the *practice* of employing incentives involving inequalities, given the inherent selfishness of persons. A society is warranted in *granting* higher salaries, given its members lack of altruism; though said members are themselves unjustified in failing to act for the 'common good'. Thus, Rawls is not compelled to jettison one or the other of his principles. Pace Narveson, his system allows for a 'tempered' inegalitarianism according to which remedial justice may come at the expense of economic liberty.

A caveat is in order here. The recruiting mandated by AA1 must be directed at the actual victims of injustice and not all and only the members of that group or groups a significant portion of whose members have been treated unfairly. The common features of the members of such groups should serve only as a means of identifying *potential* candidates for redress, with the facts of personal history ultimately deciding the issue. Moreover, we would need to look beyond these groups to find other individuals to whom AA1 (or any other just remedial measure) ought to be applied. It is the effects of injustice on one's own life, not the lives of those with whom one may be categorized, that determines the extent to which one is entitled to the provisions of any compensatory policy.

Does AA2 also comply with Rawl's principles? Behind the veil of ignorance, one would reason as follows: "If I were someone against whose interests AA2 had operated, I would not thereby have been deprived of my fair share of the primary goods. For I would not have been prevented from competing for the position I'd failed to attain. My failure would be determined by my inability to outdo another aspirant who would have outperformed me had she not been disadvantaged. (I am assuming that the individual in question is disadvantaged and not merely the member of a group most of whose members are disadvantaged, and that I myself am not a victim of injustice despite belonging to a group in which this is not the norm.) I would not have 'lost out' to someone to whom I was superior according to the operative criteria. If I had, I would have grounds for a just complaint, as, in effect, there would have been fewer coveted positions open to me than others. But that was not the case. Moreover, I would have suffered neither a loss of freedom nor of any other primary good." Thus, AA2 would also be acceptable to someone in the original position.

It should be emphasized that the scoring of candidates must be based on publicly known, objective criteria and subject to review. Ties should neither be manufactured nor ignored. The potential for abuse must be acknowledged. But it warrants safeguards and vigilance, not abandoning the policy. (The same should be said in response to the citing of the possible misuses of AA1.) Moreover, the supposed value of *diversity* in institutions plays no role in justifying either AA1 or AA2. The self-interested parties in the original position are only concerned with mitigating the effects of injustice on its *individual* victims. As with the application of AA1, cases are meant to be decided on the merits of the aspirants involved, thus forestalling the morally offensive possibility of a candidate being rejected for reasons having nothing to do with her qualifications.

By now it should be obvious why someone in the original position would reject AA3: its application would lead to inequality of opportunity. Under AA3, victims of injustice would be allowed to compete for the positions that are set aside for them as well as those open to all aspirants; other members of society would be permitted to compete only for the latter positions.

Unlike AA1, which is intended only to make all qualified persons aware of existing opportunities, AA3 involves excluding certain individuals from the competition for some of those benefits. In contrast to AA2, according to which one's disadvantaged status can serve as a tie breaker, AA3 would allow the victims of injustice to forego having to

compete against *all* aspirants to certain coveted positions. Such a remedial measure would, thus, have an unjust downside and therefore be rejected by someone in the original position.

It would be rejected even if its application was expected to comply with the second principle's provision that all inequalities lead to an improvement in the standard of living of each member of society. Certainly it is in the best interest of each member of a society that the victims of its injustice be compensated, if for no other reason than that it lessens the chance of civil unrest. And behind the veil of ignorance, one would have, as indicated above, a conclusive reason for providing for their redress. But the second principle also places a limit upon what may be done in this regard: no one is to be left with fewer opportunities than others; inequality of opportunity is to be rectified by allowing its victims to now compete for any of the positions from which they had formerly been excluded, not by granting such persons opportunities not available to others. That applying AA3 would, in effect, replicate the very inequality it is supposed to remedy, thus, warrants disregarding the fact that it would redress victims of injustice as well as improve, in some respects, the lot of everyone else. Rawls suggests as much in the following passage:

First, though, I should note that the reasons for requiring open positions are not solely, or even primarily, those of efficiency. I have not maintained that offices must be open if in fact everyone is to benefit from an arrangement. For it may be possible to improve everyone's situation by assigning certain powers and benefits to positions despite the fact that certain groups are excluded from them. Although access is restricted, perhaps these offices can still attract superior talent and encourage better performance. But the principle of open positions forbids this. It expresses the conviction that if some places were not open on a basis fair to all, those kept out would be right in feeling unjustly treated even though they benefited from the greater efforts of those who were allowed to hold them. They would be justified in their complaint not only because they were excluded from certain external rewards of office such as wealth and privilege, but because they were debarred from experiencing the realization of self which comes from a skillful and devoted exercise of social duties. They would be deprived of one of the main forms of human good (84).

In a Rawlsian society, compensatory justice would not be effected by requiring some members to sacrifice important elements of personal fulfillment, only one of which, it should be noted, is financial.

Thomas Nagel misses this later point in defending AA3. He argues that it is "probably" not unjust, since it remedies a "grave social evil" and is only a temporary deviation from a "meritocratic system" that is itself morally indefensible. (8) Moreover, the sacrifices required of its victims are not so great, given that they retain the "general social dominance" out of which a situation requiring affirmative action arose.

While I for one agree with Nagel that the economic advantages generated by meritocracy are without moral justification, the complaints of AA3's victims do not stem solely from financial concerns. As Rawls emphasizes, having fewer educational or employment opportunities also decreases the likelihood of an individual fully developing her talents or professional skills. (The white male Ph.D.s who have been shunted to the periphery of academia, e.g., can speak to the frustration entailed by such a disadvantage.) Arguing that egalitarianism is unjust in any event can only fail to comfort to those who have reason to believe that they have been unfairly deprived of something whose value can not be measured in dollars and cents. Furthermore, it is just false that most of the present day victims of AA3 occupy privileged positions or are responsible for the injustice it aims to

redress. Many have fewer advantages than some of its beneficiaries. Thus, neither of Nagel's reasons warrants accepting AA3 as even a temporary measure.

Neither does an appeal to the value of diversity, integration, or racial/ethnic/gender harmony. These social goods must not be purchased at the cost violating individual's rights. It seems a mark against consequentialist theories that they would allow us to disregard this expense in the pursuit of the "common good." Defending the Kantian framework of Rawls' theory is beyond the scope of the present work. Suffice to say that the burden of proofs seems here to lie with those who would permit inequality of opportunity (the elimination of which was the original stated goal of many civil rights advocates) in the course of securing the above goods.

### III

Are there remedial measures besides AA1 or AA2 that one could accept in the original position? As we have just seen, one therein would not opt to set aside a certain number of coveted positions for the victims of injustice. But nothing that was said against that policy would apply to remedial measures whose purpose is to insure that all aspirants to such positions are well prepared to compete for them. Behind the veil of ignorance, one would accept not only AA1 or AA2 but also any other remedial measure that would promote what Rawls calls "pure procedural justice" (86).

A society exhibits pure procedural justice in case it utilizes a system of distributing its primary goods such that the resulting allocation is fair whatever it turns out to be (87). Such a system would be consonant with the two principles of justice (88).

Rawls further notes that a key element of a fair distribution system is "education for all," which "underwrite(s)" equality of opportunity (88). A just society, he contends, would prepare its youth to meet the competitive demands of adulthood. I would add that, in the case of victims of injustice, the required education would have to address not only their academic deficiencies but also help establish the self-esteem they are likely lacking. Obviously, these measures would have to go beyond providing such persons with the educational opportunities all others enjoy, since they are unlikely to be able to take advantage of them sans a certain amount of remedialization. The victims of injustice, in other words, are owed more in the way of education than that which fair equality of opportunity would require were no such persons to exist.

That some individuals are unable or unwilling to take advantage of educational opportunities does not cancel a society's obligation to make them universally available. The fact that schooling fails to improve the prospects of some persons should not weaken our resolve to educate those who *would* thereby benefit. From the perspective of the original position, one could only arrange access to all of the opportunities of which one *might* turn out to be capable of taking advantage. One could not attempt to tailor one's options to suit a particular set of abilities. Moreover, that educational efforts are going to be wasted on a particular individual is not something that can be predicted. Each person is thus owed the opportunity to benefit, to the extent to which she is capable, from a good education.

Thus, in addition to AA1 and AA2, a Rawlsian society would compensate the victims of injustice by redoubling its efforts at providing high quality, universal primary and secondary schooling. To someone in the original position this endeavor would appear highly desirable: since therein it is not known who would be educationally disadvantaged but for its existence. It would be understood as a safeguard against historical injustices such as racism, sexism, religious bigotry, homophobia, and poverty determining anyone's share of the primary goods, leaving nature and one's efforts as the deciding factors. Rawlsian affirmative action would begin at the schoolhouse.

## Notes

\* A version of this paper was presented at the 1998 APA Pacific Division Meeting. For their helpful comments at my session, I thank the respondent, Daniel Scoccia, as well as Richard Arneson and Chris Naticchia. I am also grateful to Michael Anderheiden for the suggestions he made following my reading of this paper at the 1998 World Congress of Philosophy and to Carl Cohen, Paul Hughes, Jan Narveson, Lisa Newfield, Louis Pojman, and Robert Simon for correspondence that led to improvements in earlier drafts of this paper.

(1) *A Theory of Justice*, Cambridge: Harvard University Press, 1971. Numbers in the text refer to the pages of this work.

(2) The distinction about to be drawn is also found in Louis Pojman's "The Moral Status of Affirmative Action," *Public Affairs Quarterly*, April, 1992, 181-206. My definition of affirmative action, thus, corresponds to neither the original nor the current understanding of the practice. The latter equates it with AA3, as explicated below, while the former, as spelled out in the Civil Rights Act of 1964, mandated efforts to eliminate unfair hiring practices, rather than redress the victims thereof. Moreover, affirmative action should not be confused with the promotion of racial, ethnic, or gender, "diversity," the justification for which has nothing to do with compensating the victims of injustice.

(3) Having done my best here to defend what follows as a reasonable extension of Rawls' system, I would add that I am not wed to the idea that it be taken as such. If the connection between ideal and nonideal theory is not as I have portrayed it, if the latter is not to be circumscribed by the former, then the ensuing views on affirmative action may simply be understood as those that would follow if one were apply Rawl's method, against his own stricture, to the issue of compensatory justice.

(4) In correspondence and in "A Puzzle About Economic Justice In Rawls' Theory," *Social Theory and Practice*, vol. 4 #1, pp. 1-27.

(5) "A Puzzle About ...," p. 3.

(6) *Ibid.*, p. 7.

(7) *Ibid.*, pp. 12-13.

(8) Thomas Nagel, "The Policy of Preference," in *Mortal Questions* (London: Cambridge Univ. Press, 1979) pp.91-105.