RAWLS VERSUS UTILITARIANISM IN THE LIGHT OF *POLITICAL*

LIBERALISM

(published in *The Idea of a Political Liberalism: Essays on Rawls* (Lanham: Md: Rowman and

Littlefield, 2000)

Richard J. Arneson

The critique of utilitarianism forms a crucial subplot in the complex analysis of

social justice that John Rawls develops in his first book, *A Theory of Justice*.[1] The

weaknesses of utilitarianism indicate the need for an alternative theory, and at many

stages of the argument the test for the adequacy of the new theory that Rawls elaborates

is whether it can be demonstrated to be superior to the utilitarian rival. The account of

social justice shifts in the transition to Rawls's second great book, *Political Liberalism*.

The account of what is wrong with utilitarianism undergoes revision as well. In this

essay I examine both the initial critique of utilitarianism and its transformation in Rawls's

later writings. To anticipate my conclusion: Rawls's proposal that we should maximin

rather than maximize leads to an interesting standoff. The argument for maximin is not

compelling, but straight additive maximization of the utilitarian sort is revealed to be

merely one possible function among many, any of which (for all we know) correct

morality might instruct us to maximize. Rawls further urges that utilitarianism goes

astray in taking the maximandum, the thing to be maximized, to be utility rather than

primary social goods. The argument for primary social goods is not compelling, but it

does not follow that utility alone is to be maximized. The espousal of the ideal of

legitimacy in *Political Liberalism* does not affect these conclusions, and the arguments

advanced to support that ideal are either diversionary or question-begging with respect to the debate between utilitarianism and Rawlsian justice as fairness.

For purposes of this exercise, *utilitarianism* may be understood as the doctrine that an institutional arrangement, a social policy, or an individual action is morally right just in case it is the one that, compared to the available alternatives, maximizes utility. Rawls assumes that justice is the overriding, preeminent part of the morality of instituitions, so he formulates the doctrine in this way: "society is righly ordered, and therefore just, when its major institutions are arranged so as to achieve the greatest net balance of satisfactions summed over all the individuals belonging to it" (p. 22). He understands utility (human welfare) as the satisfaction of rational desire.

TAKING RIGHTS SERIOUSLY

Perhaps the animating philosophical idea in *A Theory of Justice* is that utilitarianism does not take rights seriously, that not taking rights seriously is a grave defect, and so we need a theory of justice that better fits our core convictions about ways that people must not be treated. Slavery is morally wrong because it violates fundamental moral rights of the persons who are enslaved. Suppression of speech intended to persuade that concerns public affairs and how people might best conduct their lives is morally wrong at least when the ground for suppression is the prospect of harm that might be caused via the audience's understanding of the speech. If the fundamental moral requirement were to maximize human welfare, then it would seem that whether or not slavery and suppression that violates the core of freedom of expression are morally wrong would depend on the outcome of complex and uncertain empirical calculation about which policies work best over the long run to achieve the utilitarian aim. Rawls

observes that our convictions regarding slavery and freedom of expression are not tentatively and uncertainly held, and not held on the basis that utilitarianism posits. Even if a utilitarian theory conjoined to a complex set of plausible empirical claims could support the positions that slavery and violation of freedom of expression are morally wrong, utilitarianism is in tension with the strength of our moral conviction in these matters and seems not to capture our intuitive grounds for these convictions. These concern individual rights. Hence we need a theory of rights, a genuine theory of justice, that makes sense of our core convictions about individual rights and extends these core judgments to controversial cases in a plausible way.

Reply: In a utilitarian moral system, individual rights, if present at all, will be derived from the single fundamental aim of utility or welfare maximization. Roughly, the idea is that the recognition and protection of individual rights better promote the utilitarian goal than alternative practices, policies, and acts. Rights function to simplify and coordinate decision making among imperfectly informed individuals of limited reasoning powers and limited altruism. If we were to eschew rights and directly apply the test of utility on each occasion of acting or implementing social policies, the results would predictably be less successful, from the standpoint of utility maximization, than the results of instituting and promoting the recognition of rights. Recognition of rights involves proclaiming their moral importance and socializing and training individuals to give extra weight to rights that impinge on their practical deliberations than to their own fallible utility calculations when these conflict with rights. In this sense a utilitarian theory can take rights seriously without assigning rights any nonderivative moral significance.

The utilitarian will also note that common-sense agreement that rights are uncontroversially decisive determinants of what we ought to do is in a sense illusory. Many rights appear uncontroversial when they are stated vaguely and at a high level of abstraction, so that the practical implications for policy of acceptance of these abstract rights are highly uncertain. Take freedom of expression, for example. Almost everyone is for free speech, but this appearance of unanimity quickly dissolves if we ask what a right to free speech is supposed to entail in a host of complex circumstances. It does not follow that there is no way to proceed except by appeal to utility as John Stuart Mill argues, but the appearance that the embrace of utilitarianism would force us to regard what are really simple and obvious moral truths as contingent and uncertain matters is misleading.

THE SEPARATENESS OF PERSONS

Rawls urges that utilitarianism "does not take seriously the distinction between persons" (p. 27). The objection is that utilitarianism extends to interpersonal conflicts of interest a rule of maximizing aggregate benefit that is morally unproblematic when what is at issue is conflict between a person's interests at earlier and later times of his life. Prudence dictates accepting a smaller pain now to avoid a larger pain at a later time, and utilitarian maximization dictates imposing on one person a smaller pain in order to avoid a larger pain for another person. But in the intrapersonal case Jones now is compensated by the gain to the same Jones later, whereas in the interpersonal case, a loss imposed on Smith is not compensated by greater benefits that accrue to other persons. This last formulation might suggest that Rawls intends to assert the principle that there should be no sacrifice imposed on one individual for the benefit of others unless the individual who

suffers imposition is compensated. But to mention an example described by Philippa Foot, we may leave a row of beautiful shade trees standing along a roadway even though we know that in the long run some individuals will be killed in a roadway crash who do not voluntarily consent to bear this risk and who would have lived but for the presence of the trees. Surely not all such imposition of risk is morally forbidden whatever the ratio of cost imposed to gains secured. It is better to regard Rawls as making the point that the utilitarian principle is indifferent to the distribution of utility across persons. Utilitarianism bids us to do whatever maximizes the aggregate of utility regardless of the different distributions of utility the alternative actions we might take would achieve. But the distribution of utility across persons is morally significant, so it is a flaw that utilitarianism would have the decision about what should be done vary only with the utility total that different acts could achieve.

PRIMARY SOCIAL GOODS NOT UTILITY

Rawls's rejection of utilitarianism is structurally complex. He embraces two views of justice, a general conception that is supposed to be valid at all times, and a special conception that is valid under modern social conditions. For now, let's focus on the general conception. It holds that the major institutions of society should be set so as to make the worst off representative person as well off as possible, with benefits measured in terms of primary social goods—goods distributable by society that a rational person would want more rather than fewer of, whatever else she wants. Rawlsian justice so conceived differs from utilitarianism in two respects: we are to maximin rather than maximize, and the basis of interpersonal comparison for social justice is primary social goods not utility. Consider the latter difference.

In later writings Rawls refines the definition of *primary social good*. For simplicity the discussion in this essay employs the refined notion. According to it, these goods are ones that would be wanted by any rational person who seeks to develop and exercise her moral powers to cooperate with others on fair terms and to develop and reflectively adjust a conception of her final ends, as well as seeking means to satisfy whatever final ends she has.

The comparison of a primary goods standard and a utility standard varies with one's conception of utility. One conception identifies utility with rational preference satisfaction—satisfaction of preferences that would withstand rational scrutiny. Another conception identifies utility with objectively valuable goods, goods that appear on an "objective list," and are valuable for a person independently of the subjective attitudes she has toward them. For my purposes the differences between these conceptions of utility will not matter. On both views, utility is not to be identified with happiness or any other quality of experience. Nor is it identified with simple desire satisfaction.

Use of a primary goods standard in a theory of justice involves a division of moral responsibility between society and individual. Society is responsible for securing fair shares of liberties and other primary goods to individuals. Given this fair background, each individual is responsible for developing a set of final ends and a plan of life to achieve them and for organizing her own life to satisfy this plan. The individual is responsible for these aspects of her life in the sense that any deficits in the outcomes she reaches do not trigger further claims for compensation by society. Society is obligated to provide fair shares of resources to individuals, not to secure them any guaranteed level of quality of life or utility. Citizens are viewed as competent to set their own ends and live

their lives as they choose. A primary goods standard also insulates citizens, who together constitute society, from the open-ended, perhaps excessive obligations to others that would be a feature of a social justice regime in which society was obligated to ensure a desirable quality of outcomes after resources are used. For example, people who have developed expensive needs are not thereby entitled to more resources than people who have chosen a life in which their needs are modest.

The theme of responsibility that Rawls's discussions link to a primary goods standard does not succeed in justifying that standard. Suppose we accept the plausible idea that a theory of social justice involves a fair division of responsibility between society and individual, which will yield the result that sometimes individuals behave in a faulty self-damaging manner for which they alone, and not society, should bear the costs, and that sometimes individuals choose in a substantially voluntary way to pursue some end other than their own well-being, and any resultant individual well-being deficit they suffer should not trigger further claims for compensation by society. This acceptance of personal responsibility is fully compatible with rejection of a primary goods standard of interpersonal comparison. We might instead link personal responsibility to a utility-based approach by holding that distributive justice requires that individuals have a fair share of opportunity for utility (well-being) over the course of their lives. To a first approximation, one has an opportunity for a good provided that, if one chooses the good, one obtains it. To develop this sketch of a principle into a workable principle, one would have to specify under what circumstances one enjoys a fair share of opportunities. One factor here will be the cost of providing an opportunity to an individual that consists in lost opportunities for others. If the cost of providing me the opportunity for an apple is

that ten persons lose the opportunity for an orchard, the cost is probably too high. Another factor is that other things being equal, if one has little opportunity for utility as measured by the highest level of utility one can reach if one chooses prudently, then providing increments of opportunity to this person who is poor in opportunities is morally more valuable than providing same-sized increments to those who are already rich in opportunities. A third factor is that the idea of fair shares should take account provision over the course of an individual's life, not just the sum of opportunities at a moment, or from now. If I have squandered my opportunities when I was young, there may be nothing unfair in the circumstance that my opportunities now that I am old are small compared to those of people who husbanded their opportunities wisely. A fourth factor is that one may forfeit one's claim to opportunities in whole or in part by morally wrongful action and especially by illegal acts chosen so that one is responsible for them.

Another factor is that one may have an opportunity for a benefit in the sense that if one chooses it, one gets it, yet one may fail to choose it either because one has substandard choosing ability or that one fails to exercise such ability as one has through no fault of one's own. An example of nonculpable failure to exercise: suppose that you and I both have the ability to choose prudently with respect to personal savings and investment decisions, but exercising this ability if fun for you and excruciatingly painful for me. For this reason opportunity as defined in the previous paragraph is not quite the notion we need to specify a fair opportunity for utility principle. The relevant issue is what level of utility one would reach if one did as well as could reasonably be expected, given one's choicemaking deficiencies.

An opportunity for utility standard evidently differs in theory and in practice from a primary goods standard. Two individuals may have different abilities, given by nature or nurture, for transforming primary goods into utility. With the identical allotment of primary goods, the two individuals will have very different opportunities for utility. One is legless, the other has two good legs; or one suffers from chronic severe headaches, one does not. Also, individuals may have different abilities to perceive the good and to select genuinely worthwhile final ends. Again, with the same bundles of primary goods, two individuals with different value setting and choice making abilities will have very different life prospects. Moreover, all of these differences that influence people's capacities to transform resources into utility can occur through no fault of the individuals who end up disadvantaged in these respects. To reiterate the main point: Acceptance of appropriate norms of personal responsibility does not per se supply any reasons for rejecting a utility-based standard of interpersonal comparison, qualified by personal responsibility, in favor of a resource-oriented conception such as a primary goods standard.

My defense of utility as a standard of interpersonal comparison for the theory of justice has not so far addressed the question of how to defend a particular standard of utility or human well-being, one that is selected nonarbitrarily from rivals, enables one to make the cardinal interpersonal comparisons that principles of justice in the utilitarian family will require, and coheres with our considered judgments in reflective equilibrium. My own belief is that the best conception of utility will not identify it with fulfillment of the individual's actual preferences, aims, or plan of life. These subjective attitudes and ends can fail to track what is truly valuable, so an objective ideal of utility is needed

(whether this would turn out to be a rational preference satisfaction theory or an

"Objective List" theory or something else is a task for another occasion).

The defender of a Rawlsian primary goods standard might interject a skeptical

response at this point: the utility-based conceptions of justice are nonstarters because no

satisfactory standard of interpersonal comparison for the purpose can be developed. In *a

Theory of Justice*, Rawls does not emphasize the difficulties about interpersonal

comparison, and indicates the superiority of justice as fairness lies along some other

dimension. The interpersonal comparison problem is no doubt significant, though in my

judgment not insoluble. Here I wish to make a more limited point: Rawls has his own

unsolved difficulties with interpersonal comparison, so, living in a glass house, he is

poorly placed to be throwing stones at utilitarian windows. According to Rawls, there are

several primary social goods. Some of these qualify as basic liberties, so are treated

within the equal liberty principle, which is accorded strict priority within Rawls's system.

But this leaves several primary goods other than basic liberties, so in order to apply the

difference principle, which requires that the expectations of these primary goods be

maximized for the worst off class of persons, we need an index of these goods: a way of

determining, for any disaparate bundles of these goods, which contain more primary

goods overall. If one bundle containing various amounts of various primary goods can

be matched with another bundle that contains more of each of these distinct primary

goods, then the second bundle dominates the first, and unambiguously contains more

primary goods overall. But for the many cases where dominance does not hold, we need

an index. I do not see how an individual's bundle of primary goods can be assessed

except in terms of the extent to which those goods enable the individual to satisfy her

preferences, or to fulfill some given objective conception of the good, and neither of these ways of assessment provides a measure that is consistent with Rawls's strategy of argument and core assumptions. To my knowledge, the only serious discussion of the index problem for primary goods is in John Roemer, *Theories of Distributive Justice*.[2] Roemer asserts that the index problem admits of a solution, but his proposed solution compromises Rawls's avoidance of utility-based measures, and causes Rawls's principles to unravel, so this is not a friendly construal of Rawls that could be used to defend his position against a utilitarian critic. Without pressing this issue further, at the very least there is a problem here that the defender of a primary goods standard would need to solve, and has not solved to date, if the primary goods approach to justice comparison issues is to be a viable position.

MAXIMIN NOT MAXIMIZE

Rawls's general conception of justice holds that the basic structure of society should be arranged so as to maximize the long-run expectation of primary goods for the group of members of society that is worst off in this respect. As noted above, one important aspect of this principle is its use of primary social goods rather than utility to measure the condition of individuals for purposes of social justice. Another crucial aspect of the doctrine is the imperative to maximin, to make the worst position in society as tolerable as possible.

A straight maximizing doctrine such as utilitarianism that bids us to arrange society so as to maximize the aggregate sum of benefits enjoyed by persons yields different implications from a maximin principle in situations in which the latter's strong tilt in favor of the worst off looks attractive. We can imagine a world in which

utilitarianism would require taking away large and important benefits from a small number of the very worst off, driving them to the wall as it were, in order to yield trivial benefits for a very large number of persons who are already enjoying an extremely hight level of benefits. We are asked to take the corn from the serfs to add one more layer of frosting to the desserts of the very rich. Here a revulsion against maximizing the aggregate sum of benefits is reasonable.

But there are cases and cases. A maximin principle is entirely insensitive to the numbers of the worst off who gain or lose, by comparison with the numbers of the other members of society who might be asked to sacrifice for the benefit of the worst off. Maximin is also utterly insensitive to the amount of benefit that is gained or lost by the worst off, by comparison with the amount that others stand to lose. This means that maximin would prefer the outcome in which a single worst off person gains a penny's worth of benefit at the cost of the loss of thousands of dollars for each of thousands of the better off. A maximin rule introduces a strict lexical priority for the interests of the worst off, however slight, when they conflict with the interests, however great, of the next worst off. Lexical priority relations among moral values are strong medicine, and perhaps are very rarely if ever justifiable if the value give lexically lower priority really has value at all. In the context of the distribution of social benefits as assessed by social justice principles, lexical priority for the worst of is implausibly extreme.

Rawls has an interesting response to this obvious objection. He asserts that maximin is not a good choice rule in general, but makes sense in the special context of choice of principles regulating the basic terms of social cooperation. And it makes sense despite yielding implausible implications in some possible cases because these

implausible cases are never actually confronted in this-worldly distributive justice conflicts. The scenario in which we can gain a penny for a single worst off person at the cost of huge losses for huge numbers of the already better off individuals is simply unrealistic. Hence maximin is acceptable as a this-worldly principle of social justice even if not acceptable as a moral principle to govern all possible worlds.[3]

This reply invites two further objections. First, it is simply untrue that plausible this-worldly scenarios cannot be imagined in which we could face a choice of very small gains for the worst off at the cost of very large losses for the better off. Consider public education policy viewed as a mechanism for enhancing the job and income prospects of low-skilled persons. We could spend more money on the least skilled and less on the more skilled, and we could continue doing this even if the further gains from further educational resources deployed on the least skilled produced very marginal increments in the labor market and entrepreneurial market prospects of this very disadvantaged group. The least natively skilled, in the absence of special education, may well be very poor transformers of educational resources into enhanced skill levels, and even if there is a range within which extra education for the worst off produces significant gains, maximin require us to keep transferring resources to the worst off until further expenditures produce no extra gains at all. At this point the curve that plots expenditures versus benefits in terms of enhanced skills might be very unfavorably steep. I do not believe that maximin provides a plausible guide to an egalitarian, socially progressive public education norm. One might object that at some point rather than moving more educational resources to the education oif the worst off, we should instead expand the education of the talented, who will then move into productive slots in the economy, and

the extra productivity their extra education generates can be partly captured for the worst off in the form of tax and transfer policy. But first of all this depends on the tradeoff in an index of primary goods between the good of extra income and the good of enhanced chances of getting hired at better jobs. The latter good is gained by increased education but not by a tax and transfer policy. But more fundamentally, a tax and transfer policy which gives lexical priority to producing gains for the worst off can itself, pushed to the limit as the principle requires it must be, will itself eventually be generating very tiny gains for the worst off that are secured by very large losses suffered by the better off.

Rawls's criticism of utilitarianism commits him to a standard of adequacy for proposed moral principles that the maximin principle on his own account of its merits cannot meet. Recall that Rawls had objected that even if it turns out that utilitarianism in the actual circumstances we face or are likely to face would imply policies of respect for individual rights that we find acceptable, utilitarianism would still be delivering the right answers for the wrong reasons. The utilitarian defense of rights such as the right to freedom of speech does not fit the actual reasons that on reflection we take to be the reasons that underlie our affirmation of such rights. The test of adequacy that utilitarianism cannot meet is agreement with our considered moral convictions including among our convictions our reasons for believing why the policies we support for actual circumstances are right. To put the point another way, utilitarianism as applied to issues of individual rights fails a test of counterfactual stability: Even if utilitarianism delivers plausible answers in actual and likely circumstances, we can imagine possible even if improbable circumstances in which we would continue to find that our reasons for

supporting free speech are fully in play even though utilitarianism would recommend the suppression of free speech.

But then by the same token the maximin principle is unacceptable as a principle of justice if it yields implications for policy in many possible situations that we would on reflection judge to be unjustifiable even if it yields policy judgments we mostly agree with in actual and likely circumstances. Here the critique of utilitarianism comes to haunt Rawls. The weapons he turns against utilitarianism, to some degree successfully, can be deployed with success against the maximin principle and perhaps other principles he favors.

THE GENERAL AND THE SPECIAL CONCEPTION

At this point it will be helpful to introduce a distinction that Rawls draws between the general and the special conception of justice. The general conception is supposed to hold true throughout human history. It specifies that institutions are just when over the long term the primary social goods holdings that accrue to the worst-off social group are as great as possible. The special conception of justice is intended to apply only to modern societies that are sufficiently wealthy so that it is possible for all citizens to exercise the basic liberties. The special conception asserts that (1) "Each person has an equal claim to a fully adequate scheme of equal basic rights and liberties," and (2) "Social and economic inequalities are to satisfy two conditions: first, they are to be attached to positions and offices open to all under conditions of fair equality of opportunity; and second, they are to be to the greatest benefit of the least advantaged members of society." [4] The "basic liberties" of the Equal Liberty Principle are given by a list of civil liberties such as the right to vote and stand for office in free elections and the

right to freedom of speech and assembly. Fair Equality of Opportunity requires that institutions be arranged so that any two persons in society with the same native talent and the same ambition should have the same prospects for success in competitions that confer above-average levels of primary social goods other than basic liberties. The second part of the second principle is known as the Difference Principle. Rawls specifies that the Equal Liberty Principle has strict lexical priority over the second principle and the the Equal Opportunity clause of the second principle has strict lexical priority over the Difference Principle. One principle has strict lexical priority over another when one must bring about the satisfaction of the first to the greatest possible extent and devote resources to implementing the second principle only insofar as doing so does not lessen at all the extent to which the first principle is satisfied.

One might view these principles as implicit in the public culture of modern democratic societies. In these societies maintaining equal civil liberties is commonly regarded as especially important and deserving special priority. We should not be willing to trade civil liberties for increased economic wealth, for example. Rawls presses this intuition to its logical limit by assigning strict lexical priority to the equal Liberty Principle. Another firm conviction in modern diverse democracies is that all citizens regardless of their race, creed, or sex or similar differences are entitled to equal protection of the laws. Again, Rawls presses this antidiscrimination norm to the limit by asserting Fair Equality of Opportunity, the requirement that nothing but native talent and individual ambition should affect one's success in competition for social positions that confer special advantages. Another, perhaps less widely held conviction is that society should take steps to offset the outcomes of competitive markets to ensure decent life

prospects for all members of society, even the least talented. Rawls presses this help-the-needy intuition again to the limit by proposing that institutions should be set so that the group that is worst off by the measure of primary goods holdings is made as well off as possible. If one shares these core social democratic convictions, one might object that Rawls carries them too far or ignores other values that should have countervailing weight in policy choice, but one cannot object that Rawls is halfhearted in his embrace of them or tepid in his espousal. He requires that these democratic egalitarian values, properly ordered, be implemented to the greatest possible extent.

This characterization invites an objection. Strict lexical priority relations that are supposed to govern the elements in Rawls's system, so that concern for equal basic liberties trumps concern for all else, and concern for Fair Equality of Opportunity trumps the Difference Principle, and according to the latter principle, achieving a benefit no matter how tiny for the worst off trumps the achievement no matter how large for any other group. All of these lexical priority relations turn out to be implausible when one contemplates them. I have already remarked on the implausibility of the maximin idea that the Difference Principle expresses. To take another example, we would not judge a society that did worse than another on the score of basic liberties even by a very slight amount to be necessarily less just overall than the society that does slightly better with respect to civil liberties no matter how dreadful its comparative record with respect to other justice values. But the lexical ordering of Rawls's principles of justice with the Equal Basic Liberty principle in first place enforces the contrary judgment.

It is not clear how seriously Rawls expects us to take his stipulated lexical priorities. He offers the idea in a tentative and provisional spirit (p. 45). But he is led to

tentative assertion of lexically ordered principles because he believes that a fully successful rival to a utilitarian theory of justice must be a genuine theory in the sense of a set of principles that by themselves determine what should be done without the need for further ad hoc moral judgment given any specification of the factual circumstances in which the principles are to be applied. J.S. Mill in his essay on *Utilitarianism* had asserted this completeness requirement as a condition of adequacy for a candidate moral theory. Mill believed that utilitarianism satisfied this adequacy condition and that this fact gives the utilitarian doctrine a leg up on its competitors.[5] One might then hold that if Rawls's lexical priority relations are deemed unacceptable, then his justice as fairness conception minus the lexical priorities does not constitute a genuine theory of justice and cannot then win the competition against utilitarianism. After all, it is said that you cannot beat a theory except with another theory.

This point does not seriously undercut Rawls's criticisms of utilitarianism. If Rawlsian arguments convince us that utilitarianism cannot do justice to our considered moral judgments after reflection, then even if we are unsure exactly what principles should be substituted for utilitarianism, uncertainty on this score does not weaken the force of the negative case against that doctrine. The position thus reached will be unstable pending the elaboration of a set of principles that match our considered judgments, and the possibility remains that if no such superior set of principles is ever forthcoming, our initial confidence that the negative case against utilitarianism is compelling might have to be revisited. But we can know that a doctrine must be wrong without yet perceiving a replacement doctrine that is corect.

ORIGINAL POSITION ARGUMENTS

The critique of utilitarianism and the defense of the conception of justice Rawls calls "justice as fairness" are developed on two separate fronts. On the one side, Rawls presents considerations that show the intuitive plausibility of the principles he favors over the utilitarian rival he rejects. The principles of justice are presented as plausible in themselves and as recommending policies that after reflection we will come to embrace. On another front, Rawls argues that his favored principles of justice would be chosen over alternatives that prominently include utilitarianism in a special choice situation that is supposed to collect and express our judgments about fair procedures for choosing moral principles. This is the celebrated and vexed original position argument.

The core idea of the original position is that parties choosing principles that are then imagined to regulate the society they will inhabit and choosing on a basis of self-interest under ignorance of all particular facts about themselves (they are stipulated to know only general facts about the world given by physical and social science) would choose Rawlsian principles of justice. Rawls mainly argues that the parties in his original position would choose his favored principles rather than a version of utilitarianism that calls for maximizing the average level of utility in society over the long run. The original position in Rawls's hands is a complex construction. To some extent it is designed with a view to producing as output principles we find intuitively plausible, and to this extent the device showcases conclusions already reached and does not involve independent argument for those conclusions. But to some extent Rawls believes that the choice situation he constructs collects intuitions about fair procedures for choice of principles and fair procedures for securing agreement that have justifying force independent of our

preexisting moral judgments about the principles that might emerge as the outcome of the procedure.

Here I want to register dissent. I do not believe that we have any moral intuitions about fair procedures at the level of abstraction at which Rawls is working. We do not have any ideas about fair procedures for choosing moral principles. The best procedure is whatever is most likely to lead to the choice of principles that make sense. Rawls suggests that it is appropriate to place the agents in the original position under a thick veil of ignorance which deprives them of information that might be used to bias the choice of principles. But this claim gets things back to front. Since we have antecedent strong opinions that moral principles should not favor individuals of any particular sex, race, or religion, we do not want the original position structured in such a way as to render it likely that principles expressing bias of this sort would be chosen. Any weak views we might have about how the artificial choice situation of the original position should be designed are entirely determined by strong preexisting views about the appropriate content of the principles to be chosen. I would say that in general our notions of fair procedures are subordinate to our convictions about the outcomes that we think it fair for the procedures to reach—fair procedures in general are those most likely reliably to lead to outcomes that are fair. Whether this is correct or incorrect, it seems hard to deny that we do not have intuitions about procedural fairness when the procedure we are talking about is supposed to be picking fundamental moral principles to serve as a standard of social justice.

Despite my skepticism about the original position project, it is worthwhile looking briefly at the arguments Rawls deploys there against utilitarianism, for they may shed

light on the character of his opposition to utilitarianism and help us decide to what extent Rawls's criticisms of that doctrine suffice to exhibit his favored principles as morally compelling.

Rawls stipulates that the parties in the original position, knowing no particular facts about themselves including their desires and aims, are motivated to gain primary social goods for themselves. So the issue that is salient in the original position is not primary social goods versus utility, but rather maximin versus maximize. If the parties prefer more primary social goods to fewer, and know that once the veil of ignorance is lifted they could be anyone in society in any social position, one might suppose that they would choose a principle that bids society to maximize the level of primary social goods per person over the long run. Maximizing the average level would be the best one could do to maximize one's own individual level if one were entirely ignorant of the person one might turn out to be in the society to be regulated by the principles one is choosing. This line of argument for maximizing the average is what Rawls sets himself to rebut.

At this point we might stand back and ask what power this derivation of maximizing-the-average would have to justify a view like average utilitarianism in the face of objections that maximizing the sum of benefits ignores our moral concern for fair distribution of benefits. Suppose it were the case, contrary to Rawls's intention, that the original position he designed relentlessly yielded the conclusion that the parties would pick a principle of maximizing-the-average. What then? I would say this result would have no justifying power whatsoever against the initial pre-original-position claim that a singleminded focus on the average benefit level incorrectly leaves out of account the genuine moral significance of fair distribution. These intuitions about fair distribution

(just by themselves) would lead us to embrace some distribution-sensitive welfarist consequentialism rather than the average utilitarian variant of welfarism. Being told that self-interested parties choosing under extreme ignorance would pay no heed to issues of fair distribution and would choose a principle of maximizing-the-average to regulate a society they would inhabit should not cause us to doubt our beliefs in the significance of fair distribution for a moment. After all, the parties in the original position are given no concerns that would make them pay heed to fair distribution. Their ignorance is not ours.

If despite these objections we proceed further into the original position labyrinth, we find that Rawls is concerned to remove the air of implausibility that attaches to the lexical priority aspect of the maximin principle. This principle asserts that we should at all costs bring about the best possible outcome for the very worst off. As we have already indicated, and as Rawls is fully aware, this means that we should prefer to gain an extra penny for the worst off rather than to gain instead any amount of benefit however large for any number of better off persons. Rawls says that maximin is not in general an acceptable principle of rational choice, but becomes that when three conditions are met, and these three conditions are satisfied to a high degree in the special circumstances of the original position. The conditions are: (1) knowledge of probabilities is unavailable, so one cannot estimate the chances of ending up in the worst-off position or any other. (2) "[T]he person choosing has a conception of the good such that he cares very little, if anything, for what he might gain above the minimum stipend that he can, in fact, be sure of by following the maximin rule" (p. 154). (3) The principles competing with maximin may produce outcomes that one finds unacceptable.

The second condition seems straightforwardly not to obtain in the original position if the general facts about nature and society to which the parties have access include facts about the normal range of human motivation. In fact humans in their every day behavior are not extremely risk averse, but normally make such choices as, for example, to drive across town, risking slightly a violent death in a car accident, in order to get a fancy loaf of bread, when one could have obtained less fancy bread with no such risk of death at a store near one's house. People generally care a lot about being able to obtain ambitious life goals that require large amounts of resources to fulfill. There is an unclarity here, because in the original position, not knowing what society one belongs to or the level of economic development of that society, one cannot estimate what level of primary goods the adoption of the maximin rule would guarantee in the actual society one would inhabit. But even if we assume the level of economic development and wealth is high, the point remains that humans as we know them do attach considerable value to obtaining very large holdings of resources that are far above what maximin could be thought likely to guarantee.

The third point taken by itself does not seem to motivate choice of maximin but rather soemething like a disaster avoidance principle that would have society organize itself so as to get as large a proportion of its population as possible over some threshold level of ersource provision deemed to be necessary for a decent existence. This disaster avoidance rule would require transferring resources from worse off groups in society to the best off when only the best off can be raised over the threshold of minimal decency. Whatever the attraction or lack of attraction of such a principle, it is not to be confused with maximin.

The first point, the unavailability of information about the probability that one will end up in one or another social position, is simply stipulated to hold in the original position. We must then ask what the justification is for that stipulation. Suppose that the choice of maximizing-the-average benefit level in certain circumstances might entail a social order in which a tiny number of people lead lives of excruciating misery. If in the original position we knew that the chances of becoming one of these unfortunates was extremely small, we might find the prospect of ending up in a better social position sufficiently attractive to render choice of the average rule a good bet despite the posibility of a disastrous outcome for oneself. Ruling out probability information in the original position rules out this sort of reasoning. If one had an antecedent conviction that making the condition of the least advantaged person in society as advantageous as possible was the paramount consideration for social justice, and that no amount of benefit, however large, to any number of other people, however huge, could overbalance this consideration, we could justify the Rawlsian stipulation regarding lack of information about probabilities. But so understood, the acceptability of excluding such probability information assumes a prior conviction of the moral correctness of the maximin rule as a principle of social justice, and so does not offer independent support for maximin. And by the same token, our antecedent conviction that maximin is an extreme rule that does not permit balancing and trade-offs of benefits and losses among worse-off and better-off within a distribution-sensitive scheme that we on reflection would endorse is not impugned in the slightest by the abstract reasoning of the parties in the hypothetical original position as designed by Rawls.

POLITICAL LIBERALISM

Beginning about 1985, Rawls reconceives the project of constructing a theory of justice.[6] He tells his readers that his starting point is recognition that his theory as originally presented fails to solve the problem of stability. The problem emerges if we imagine that the institutions of a society are made to conform fully to Rawlsian principles of justice. Now Rawls notes the fact of pluralism, that in the absence of unjustifiable state tyranny, reasonable citizens of a society under modern conditions will continue to embrace different and opposed general and comprehensive ethical doctrines. This pluralism is a permanent feature of modern society, not a temporary phenomenon. An ethical doctrine is fully comprehensive when it embraces all values and virtues, and general when it applies to a wide range of subjects. Many religions aspire to be general and comprehensive, as does the utilitarian moral philosophy. Given pluralism, the citizens of a society that fully implements Rawlsian justice would not tend over time to accept its justification, Rawls's partially comprehensive theory of justice. This means that an ideal of legitimacy is violated that holds that the uses of state coercion in a community should be in accordance with moral principles that each member of the community, if she were reasonable, would accept. A satisfactory theory of justice fulfills legitimacy, so Rawls concludes that his doctrine of justice as fairness is unsatisfactory and stands in need of revision.

At this point it might seem that Rawls has painted himself into a corner. If the conditions of modern society, absent unjust tyranny, spawn continuous and unending reasonable disagreement about morals, and if an adequate theory of justice must be capable of being implemented so as to attract unanimous assent from all reasonable persons affected by it, then it looks as though an adequate theory of justice is impossible.

Rawls has an ingenious solution. What is needed is a political conception of justice, which does not aim to be a comprehensive doctrine of any sort, but rather a doctrine of fair terms of cooperation to regulate major social institutions that is designed to bypass intractable ethical, religious, and philosophical controversies. A political conception of justice suitable for a constitutional democracy with stable pluralism of belief among its citizens seeks to define fair terms of cooperation for the circumstances we face, including the fact of pluralism itself. A successful political conception will be such that it can attract an overlapping consensus of the reasonable comprehensive views held by the members of the society. In this scenario, all citizens qua reasonable agree to the basic terms of social cooperation, though they do so from different and opposed moral standpoints.

That an overlapping consensus of this sort is possible for us, Rawls holds, reflects favorable contingencies. The history of constitutional democracy shows that its basic features can attract wide support. A political conception of justice draws upon common-sense ideals that are latent in the public culture of constitutional democracy, and develops these common ideals into a system that can order and guide our judgments and that can be acceptable to all after due reflection.

By a happy coincidence, it turns out that the conception of justice that Rawls espoused in *A Theory of Justice*, though not designed to be a political conception of justice that can become the focus of an overlapping consensus, is in fact suitable to this role. In justice as fairness, a solution to the problem of devising fair terms of cooperation is solved by representing persons as fair and equal and determining what such persons would choose as fair terms of cooperation in an original position. *Political Liberalism*

adds a twist to this account. The ideal of a free and equal rational and reasonable person is now to be embraced by persons as a political ideal., a way they wish to be conceived for the purposes of settling basic rights in a diverse democracy. It is supposed to be fully compatible with embracing this Kantian ideal of the person as a political ideal that in private life one regards oneself quite differently. As a private member of society, I may regard myself as forming my beliefs by deferring to religious authority and as unable to imagine myself as renouncing my core beliefs. This is compatible with also holding that as a citizen, I regard myself as free to choose new beliefs and values, and I hold that if I were to renounce my present faith, my rights and responsibilities set by social justice norms should remain the same.

Rawls has an attractive answer to the concern that in a society divided by permanent disagreement about what ultimately matters, purported norms of justice will express the values of the citizens who happen to gain political power, and will amount to an imposition of some people's values on others who reject them. A related concern is that only a skeptic can be a liberal with respect to issues of toleration. Someone who really believes that salvation is all-important and can be gained only through his church, or that saving the planet for the benefit of all beings who inhabit it is an absolute duty and requires an ascetic lifestyle, cannot without hypocrisy assent to a political regime in which people are free to behave in ways that promote damnation or planetary ruin. The answers to both concerns proceed from acceptance of a legitimacy ideal. In politics, the legitimacy ideal requires that the coercive use of state power should not be deployed on individuals on the basis of moral principles that they reasonably reject. In the conduct of life, the legitimacy ideal requires that one act toward others only on the basis of moral

principles that they cannot reasonably reject. The legitimacy ideal can be regarded as an expression of a Kantian norm of respect for persons.

An argument against basing public policy of a diverse society on utiilitarianism or any principle in the family of utilitarian views waits just around the corner. Like any comprehensive moral or religious doctrine, utilitarianism will inevitably be controversial. Many persons will reasonably reject it. Hence any attempt to advance utilitarianism as a political conception of justice would be sectarian; it would amount to imposing a set of ultimate values on people who reasonably reject it. Basing state policy on utilitarianism or entrenching utilitarianism into the constitution of a democratic society as the foundational value would be akin to establishment of religion. No more than Buddhism, Roman Catholicism, or any other comprehensive doctrine should utilitarianism be advanced for the role of public philosophy, political conception of justice.

One signal that something is amiss in the overlapping consensus story that Rawls tells is his account of the problem of stability. He envisages moral argument aimed at finding the best account of social justice that is responsive to the problem of arriving at fair terms of cooperation among persons with diverse beliefs. The next stage of the argument is to try to determine as best we can whether this candidate conception of justice could be implemented in a modern society so as to attract over time the unanimous or nearly unanimous agreement of reasonable members of society living under this regime. This is the legitimacy ideal. If it cannot be met, this shows the candidate conception of justice is unsatisfactory and must be rethought. A satisfactory political conception of justice must be capable of becoming the focus of an overlapping

consensus including all of the reasonable comprehensive doctrines with adherents in the society.

One might suppose that if the candidate conception ex hypothesi is the one that after full critical reflection is best supported by moral reasons, then reasonable members of society must support it. If they did not, they would show themselves to be unreasonable after all. So it is hard to see how stage two of Rawls's argument, addressing the supposed problem of stability, can be other than irrelevant.

It turns out that the "best reasons" that hold sway at stage one of the argument need not be the same "best reasons" that hold sway at stage two. For one thing, the overlapping consensus that Rawls deems necessary is intended to sweep broadly and include the major traditions of ethical thought that attract allegiance in modern democracies. To this end the notion of the "reasonable" pitches its requirements at a modest level. A reasonable individual in this modest sense has a disposition to justice, an effective desire to cooperate with others on fair terms acceptable to all. The reasonable individual is specified as also having the capacity to form and act on a conception of the good. The reasonable individual is further specified as one who accepts the fact that fully reasonable people (reasonable by this same conception) my fail to agree on substantive moral principles, given what Rawls calls the "burdens of judgment" (factors that may lead individuals to hold stable conflicting views but that do not tend to impugn the rationality of any of the parties to these disagreements). Rawls finally stipulates that if a reasonable person embraces a comprehensive doctrine, this will be a reasonable doctrine, that is, "one that covers the major religious, philosophical, and moral aspects of human life in a more or less consistent and coherent manner."[7]

By now it should be evident that the *Political Liberalism* argument misfires. To recapitulate: It is proposed that a legitimate government is one that acts only on the basis of principles that all reasonable citizens accept. A legitimate government cannot then be one that imposes a state religion on all citizens even though some citizens reasonably reject the state religion. But then any government that affirms any comprehensive moral doctrine as the basis of its policies cannot be legitimate, because some reasonable citizens will reject any comprehensive doctrine such as utilitarianism. Add to this the premise that no society can be just whose government is illegitimate, and we have a quick argument to the effect that the norm that forbids liberal governments from enforcing a state religion equally forbids the enforcement of any secular comprehensive moral doctrine such as utilitarianism. If advanced as a fundamental principle of justice, utilitarianism is irredeemably sectarian. Call this the *argument from legitimacy*.

But this argument ushers utilitarianism out the front door but allows entry at the back door. Even if the argument from legitimacy were sound, all that it would rule out would be the adoption of the comprehensive moral doctrine of utilitarianism as the fundamental principle of justice. But for all that has been said so far, we could propose a simple noncomprehensive version of utilitarianism as a candidate political conception of justice. We might assert that given the intractability of disagreement about morals, a society can do no better than to adopt policies that count welfare improvements to any citizen as equally valuable and strive to maximize the sum of the welfare of all citizens. I doubt that this repackaging of utilitarianism as a stripped-down political conception of justice would succeed, so I shall not follow this line further. Since Rawls has set the requirements for qualifying as "reasonable" very low, the constraint that a legitimate

social order must be acceptable to all reasonable persons becomes correspondingly constraining, to the point that one wonders whether any very substantive morality—utilitarianism or any other--could satisfy the constraint. It surely becomes immediately a doubtful proposition that Rawls's justice as fairness could pass this very demanding legitimacy test.

More important is this: Rawls deliberately sets the requirement of "reasonable" at a low level, so that one can be fully reasonable in Rawls's sense when embracing moral doctrines on the basis of errors in reasoning or other cognitive failures. But then acceptance of the legitimacy ideal carries a moral cost. There are the fully reasonable doctrines that would be embraced as acceptable options by persons after ideal critical deliberation while making no cognitive errors. There is the different set of moral doctrines that would be embraced by all persons in a society who are "reasonable' in the special undemanding Rawlsian sense. But why is it morally better for society to be governed by the latter moral views of justice rather than the former? The views of justice acceptable to all, reasonable and unreasonable, are likely to be very thin gruel. At least there are aspects of justice that would be affirmed by all fully reasonable persons that would be rejected by some Rawlsian-reasonable persons. We can choose between achieving legitimacy to a greater degree or justice to a greater degree. I see no reason not to opt unequivocally for justice given the way that the legitimacy ideal has been constructed.

If there are competing visions of social justice in a society, then if one assumes that any one can be fully implemented, it is morally better that the one be implemented that is superior as a theory of justice, best supported by the balance of moral reasons all

things considered. If more than one conception ties for best, then one of the best conceptions should be implemented.

The idea that the same principle of state legitimacy that rules out the imposition of an established religion also rules out the organization of society according to a secular moral doctrine such as utilitarianism turns out to be a nonstarter. The legitimacy ideal should be rejected because it is warped by the Rawlsian insistence on a least common denominator notion of what qualifies a person as "reasonable" for purposes of deciding when one's nonacceptance of a proposed norm of social justice is reasonable or not.

This leaves us back where we started before the introduction of the special arguments of Political Liberalism. None of what I have said in this connection undermines the sound criticism of utilitarianism that we have discovered in Rawls's earlier writings, but the issue remains, what theory of justice is best supported by the balance of moral reasons, rather than what theory of justice could command unanimous assent among well-meaning but variously reasonable and confused and ignorant members of society.

THE EPISTEMIC ASYMMETRY BETWEEN THE GOOD AND THE RIGHT

A deep feature of Rawls's views that shapes his thoroughgoing rejection of utilitarianism is his conviction that "principles of justice are objective and interpersonally recognizable in a way that conceptions of the good are not."[8] This conviction leads him to sweep aside not only utilitarianism but any view that takes the fundamental idea of justice to be achieving a fair distribution of opportunity for genuinely good, choiceworthy lives for people. In Political Liberalism the focus of Rawls's skepticism on this point shifts, and targets comprehensive conceptions of morality rather than specifically

conceptions of human good. But in the end Rawls does believe that despite pluralism of belief about morals, we can in the end, if we are reasonable, agree on what is fair, but not on what is good and worthwhile in human life.

The method of reflective equilibrium that Rawls recommends for determining what ethical claims are acceptable does not suggest a reason for supposing that reasoned agreement about the good cannot form part of the moral consensus of a just society. Rawls in effect proposes a coherence test for the acceptability of ethical claims. The claims we should accept are those that we are disposed to believe after thorough critical scrutiny of pertinent arguments with full information, and we can imagine an ideally extended version of this process and conceive of rational ethical claims as those we would affirm in ideal reflective equilibrium. No doubt we now are very far from the position of ideal reflective equilibrium, so our assertions about ethics must for now be tentative. But this is as true of our judgments about what's fair and just as of our judgments about's what good and choiceworthy. No doubt also if we glance around a modern democratic society we will notice all manner of disagreement about fundamental ethical matters. But this disagreement spans disagreement about what's fair as much as disagreement about what's good. Moreover, some of this disagreement about the right and the good alike is rooted in ignorance, prejudice, confused reasoning, and other forms of inattention to reasons and incapacity to grasp them. The class of ethical claims that are in fact controverted is larger than the class of claims that are in fact controversial. To reiterate: in our present primitive state of development of ethical science, we have no reason to accept an epistemic asymmetry between the good and the right, much less to make such an assumed asymmetry a core doctrine of the theory of justice.

In passing, we should note that the idea that rational agents in communication with each other can end up in stable disagreement about what's true seems likely to prove incoherent under examination. If you believe that fish is good food and the difference principle is unfair and I believe initially that fish is not good food and the difference principle is fair, then when we share our reasons, the alternatives would seem to be that we rationally agree or we fail fully to appreciate the force of the reasons in play (we are not then fully rational agents). How could I continue to believe that my beliefs on the good and the right are uniquely correct in the face of your continued disbelief and my inability to articulate reasons that compel belief by both of us? If we say your experiences have led you to one conclusion and me to another, but we cannot extract any stateable reasons from our experience, how can I think my experience has epistemic significance? How can my experience rationalize my claims when I see the epistemic situation is symmetric and that you can make a parallel claim? Although this issue needs more discussion, my sense is that I must retract my claim to knowledge in the face of your claim, and you must do the same. If you have reason to believe fish is good food and I have comparable reasons to affirm beef as choiceworthy, the tentative conclusion should be that there are equally cogent considerations to find fish and beef choiceworthy, so perhaps either is good.[9]

Rawls raises the question why it is that the free institutions of a liberal society lead to reasonable pluralism of belief as opposed to reasonable agreement in belief. In ethical matters and the conduct of life we experience pluralism, whereas, says Rawls, we find reasonable agreement in the domain of natural science.

This contrast between reasonable pluralism in ethics and reasonable agreement in science is ill-posed. Scientific agreement is agreement among a rather small cadre of trained experts. Within the general population of a modern democracy all manner of bizarre and superstitious and metaphysical beliefs about empirical matters flourish. Throughout most of human history even this agreement about broad theoretical empirical claims did not exist even among a small elite of trained experts. If a modern democracy bases its policies on scientific views about the causal structure of the world, society is not relying on an uncontroversial consensus, but is rather imposing expert views on a largely unbelieving lay public. This practice does not violate the Kantian injunction against imposing on people on the basis of claims they could reasonably reject. It is not treating someone disrespectfully to treat him according to standards that he does not actually accept but would if he were fully rational. This holds true whether the standards in question are empirical or normative.

Rawls interprets the Kantian requirement concerning reasonable rejectability in an overly strict way. He regards it as a necessary condition for an acceptable theory of justice that it should become the object of a consensus that includes all reasonable members of society. According to Rawls, the concept of a <u>reasonable person</u> that is appropriate for the theory of justice should aim to be inclusive. A reasonable person according to Rawls wants to cooperate with others on fair terms and has a capacity for a conception of the good. But then a reasonable person's conception of the good can be based on confusions of reasoning or ignorance of fact. The reasonable person's conception of the good might be based on adding two plus two and getting the answer: five.

My surmise about what has gone wrong here is that Rawls wants the idea of the reasonable person to be tightly linked to the doctrine of toleration. I'm skeptical about the link. I have little doubt that many religious doctrines, maybe most, perhaps all, attract adherents on unreasonable grounds, and that significantly better grounds for the doctrines are not to be found. The reason for tolerating, say, fundamentalist Christian belief, is that interference is likely to do more harm than good. This tends to be so, among other reasons, because any popular sect's unreasonable doctrines tend to have few bad consequences in practice, since most religions, however unreasoned their basic tenets, counsel prudence, not imprudence, in the affairs of life. One should also mention the general case for freedom of speech—that toleration even of manifestly unreasonable speech creates a freewheeling context of inquiry that is more favorable to social progress in the long run.

If the set of reasonable beliefs about the good is restricted appropriately to the beliefs that are well supported by reasons, the extent of intractable disagreement about the good lessens, and the reasonable pluralism of belief that remains does not preclude the aspiration to construct a consensus on a substantial doctrine of the good that will partly determine the content of a reasonable doctrine of justice for modern democratic society.

CONCLUSION

This review of Rawls's critique of utilitarianism has reached a mixed verdict. Rawls's opposition to the idea that justice consists in maximizing aggregate human well-being is sound, but his rebound to the maximin ideal is gives excessive priority to gains for the very worst off and does not allow for the incorporation of sensible ideas about

personal responsibility into a distribution-sensitive consequentialism. His suggestion that

justice should be concerned with the distribution of resources, the primary social goods,

rather than with utility (individual well-being), strikes me as incorrect. The themes of

pluralism and diversity of belief in modern democracies that have preoccupied Rawls in

his recent writings on justice do nothing to alter these verdicts. The issue how the theory

of social justice should as it were split the difference between J. S. Mill and Immanuel

Kant remains wide open.

---

[1] . John Rawls, <u>A Theory of Justice</u> (Cambridge: Harvard University Press, 1971). Further pages references to this work are given in parentheses in the text.

[2] . John E. Roemer, <u>Theories of Distributive Justice</u> (Cambridge: Harvard University Press, 1996), pp. 165-172, and the references there cited.

[3] . For interesting further defense of maximin, see Joshua Cohen, "Democratic Equality," <u>Ethics,</u> vol. 99. No. 4 (July, 1989), pp. 727-751.

[4] . Quoted from John Rawls, <u>Political Liberalism</u> (New York: Columbia University Press, 1993), p. 6.

[5] . J. S. Mill, <u>Utilitarianism,</u> in <u>Collected Works</u>, vol. 10, ed. J. M. Robson (Toronto: University of Toronto Press, 1969), pp. 206-207. (This work was originally published in 1861).

[6] . See John Rawls, "Justice as Fairness: Political not Metaphysical," <u>Philosophy and Public Affairs</u>, vol. 14, no. 3 (1985), pp. 223-251.

[7] . Rawls, <u>Political Liberalism</u>, p. 59.

[8] . Quoted from Thomas Nagel, "Rawls on Justice," <u>Philosophical Review</u> (1972).

[9] . See Charles Larmore, <u>The Morals of Modernity</u> (Cambridge: Cambridge University Press, 1995), for an opposed view.