

John Rawls (1921-2002)

John Rawls was the most important political philosopher of the latter half of the 20th century. His major work, *A Theory of Justice* (1971), gave a new impetus to the subject, providing a fresh approach which transformed familiar debates. It is still the dominant starting point for arguments within contemporary political philosophy.

1. Life

Rawls grew up in Baltimore and entered Princeton University in 1939; after graduating in 1943 he served in the US Army, and he was in the Pacific in August 1945 when a nuclear bomb was dropped on Hiroshima, an act he later condemned as a great wrong.¹ After the war he returned to Princeton and completed a Ph.D. in philosophy in 1950. In 1952 he was awarded a Fulbright scholarship which enabled him to spend the year 1952-3 in Oxford. He returned to the USA to an academic position at Cornell, where he taught from 1953 until 1959. The Cornell Philosophy Department was then strongly influenced by Wittgenstein's work, and Rawls's writings from this period show this influence. In 1959 Rawls moved to Harvard, first just for one year, and then, after a two year spell at MIT, for the rest of his career. The dominant philosopher at Harvard at this time was W. V. Quine, whose influence can also be seen in Rawls's writings. The publication of *A Theory of Justice* in 1971² brought him both fame and controversy; and his work thereafter was largely taken up with defending, refining, revising and extending the position he had there advanced. He was appointed James Bryant Conant University Professor at Harvard in 1979. He retired in 1991, but continued to teach until incapacitated by a stroke in 1995.

Despite his fame Rawls was an exceptionally modest individual. He shunned both public debates and public honours (I once arranged for him to be offered an honorary degree by Cambridge University, but to the consternation of the University authorities he declined the offer). His wish to avoid special treatment is exemplified by the following story: '... it somehow came up in conversation with an administrator at Harvard that he was a vegetarian. "But you've gone to all those dinners without ever telling anyone," she exclaimed. Rawls's response was that he was not that interested in food and preferred not to make a fuss – he

simply left the meat on the plate'³. His general attitude to life is, I think, captured in the following passage:

It is a mistake to believe that a just and good society must wait upon a high material standard of life. What men want is meaningful work in free association with others ... To achieve this state of things great wealth is not necessary. In fact, beyond some point it is more likely to be a positive hindrance, a meaningless distraction at best if not a temptation to indulgence and emptiness. (*TJ*⁴ 257-8⁵)

2. Justice, fairness and reciprocity

'Justice', says Rawls, 'is the first virtue of social institutions' (*TJ* 3: 3). So, for Rawls, a theory of justice is a theory of *social* justice, and Rawls explains that the kind of social institution he has in mind is a society, conceived as a 'self-sufficient association of persons' who constitute 'a cooperative association for mutual advantage' (*TJ* 4: 4). Rawls's societies are, however, not voluntary associations. Membership of a Rawlsian society is involuntary: as he puts it in his second major work, *Political Liberalism*, 'we enter only by birth and exit only by death' (*PL* 135-6). Although it is sometimes possible for people to change their society by emigration, it is an important constraint on a theory of justice that this possibility should not be thought of as a live option. The rules of Rawls's societies are taken to define basic moral relationships between members, rights and duties, which define 'the basic structure of society' (*TJ* 7: 6). Rawls takes it that societies are responsible for protecting these rights and thus have legitimate coercive power necessary to fulfil this responsibility (*TJ* 240: 211). Rawls's societies are, therefore, *political* societies or, as we would normally put it, *states* - especially given his assumption that they occupy 'a definite geographical territory' (*TJ* 126: 109). Rawls, however, makes almost no overt use of the concept of a state because he holds that the traditional conception of a state's sovereignty has no legitimate application.

Justice is not simply the first virtue of political societies, it is also, Rawls says, 'the first moral virtue' (*JR* 208). What brings morality into the story is the way in which moral considerations enter into his account of the 'principles of justice':

the principles of justice may be thought of as arising once the constraints of having a morality are imposed upon rational and mutually self-interested parties who are related and situated in a special way. (*JF* 63)

This constraint of ‘having a morality’ does not imply adherence to some specific moral code. Instead Rawls has a more formal value in mind, which he initially calls ‘fairness’, and this gives him the phrase ‘Justice as Fairness’ which is the title of his most famous early paper. But its significance needs some elucidation. ‘Justice *as* fairness’ is not an identity, the thesis that justice *is* fairness (*TJ* 12-3: 11); instead Rawls’s thought is that the principles of justice are to be understood as principles whose adoption by those who are to be bound by them is the outcome of a fair procedure:

The question of fairness arises when free persons, who have no authority over one another, are engaging in a joint activity and among themselves settling or acknowledging the rules which define it and which determine the respective shares in its benefits and burdens. A practice will strike the parties as fair if none feels that, by participating in it, they or any of the others are taken advantage of, or forced to give in to claims which they do not regard as legitimate. (*JF* p. 59)

As this passage indicates, according to Rawls it is fundamental to justice that its requirements be ones which those who are to be bound by them can regard as collectively self-imposed through a procedure which is fair because the choice of the principles which define these requirements is arrived at in a way which is open, inclusive, non-coercive, and impartial. In *A Theory of Justice* Rawls describes these aspects of a fair procedure as ‘The Formal Constraints of the Concept of Right’ which apply to ‘the choice of all ethical principles and not only for those of justice’ (*TJ* 130: 112).

Around 1971 Rawls modified his position, placing fairness on a level with justice (the difference between them being that fairness applies to voluntary institutions and justice to involuntary ones) and arguing that what is fundamental to both is ‘reciprocity’, the mutual recognition by persons of each other as free and equal. In part this is just a change in terminology; compare the following passage from ‘Justice as Reciprocity’ with the second sentence quoted above from ‘Justice as Fairness’:

A practice will strike the parties as conforming to the notion of reciprocity if none feels that, by participating in it, he or any of the others are taken advantage of or forced to give in to claims which they do not regard as legitimate. (*JR* 208)

As we shall see, however, reciprocity enters into Rawls’s theory in other ways, so there is also some substance to this change. But what now needs elucidation is the way in which this value is to be understood, in particular the conception of freedom involved in the recognition

of persons as 'free and equal'. In the case of procedures for determining fair rules for voluntary institutions one can understand how the procedures should respect the status of members who have the freedom to join or leave. But once an institution is involuntary, as political society is, there is no free act whereby one joins and, equally, no freedom to leave. So what is the freedom whose recognition is for Rawls fundamental to the justice of political society and moral life in general?

3. The original position and Kantian constructivism

The answer to this question about freedom requires an examination of the way in which Rawls uses the idea of a social contract. Rawls maintains that the requirements of justice are determined by principles which the members of the society would agree to impose upon themselves in a hypothetical situation which respect their recognition of each other as free and equal. In *A Theory of Justice* Rawls famously fills out this hypothetical 'original condition', as he calls it, by having us imagine ourselves placed behind a 'veil of ignorance' which prevents us from knowing anything distinctive about ourselves, our situation and our society and then inviting us to select fundamental rules for a society in which we are to live. One might think that our ignorance would leave us quite unable to make a rational choice. But Rawls holds that we are to imagine ourselves as having unlimited knowledge of general truths about the world and also knowledge of the 'primary social goods' of human life, which he takes to be rights, opportunities, wealth and self-respect. Furthermore we are to assume that our motivations in the original position are those of people who are 'rational and mutually disinterested' (*TJ* 13: 12), by which he means that they act to promote their own self-interest without any interest in one another's interests. In the light of these positive characterisations of the original position, he argues, it is possible to demonstrate that there are two fundamental principles for social cooperation which it would be rational to adopt. First, that each person has a right to the most extensive scheme of basic liberties compatible with everyone else having a right to similar liberties; second, that social and economic inequalities are to be arranged in such a way that positions of power and status are open to all and differences in wealth and income arise only through distributive systems which work to the advantage of everyone, and especially to the advantage of the least well-off members of society (*TJ* 60-1; 53).

These principles are obviously contentious and I will say a good deal more about them. But my first concern here is with the way in which Rawls characterised the original position. His idea was that by combining the veil of ignorance with the assumption of self-interested rationality he could induce the rational choice of principles which promote anyone's interests. This was a clever idea, but an immediate source of unease about it is that it undermines the aim of presenting the principles of justice as the agreed outcome of a procedure which respects the value of reciprocity. For because the veil of ignorance obscures all the differences between people, there is no question of imagining people in the original position coming to an *agreement* concerning the principles which it would be rational to adopt. Instead, the standpoint is just that of 'one person selected at random' (*TJ* 139: 120) whose rational judgment is bound to be the same as that of anyone else. So the original position with its veil of ignorance is not a model of an ideal social contract; instead, it models the role of an ideal legislator. But that role brings with it a very different conception of justice and morality, founded not on the fundamental value of reciprocity but on that of impartial sympathy.

This objection invites the question why Rawls does not take it that reciprocity is one of the primary social goods which the parties to the original position are assumed to take into account; the answer to this is that Rawls believed that by means of his account of the original position he represented the value of reciprocity, and thereby ensured that the principles of justice thus chosen would express this value (see *PL* 305-6). As John Harsanyi demonstrated, however, this belief is mistaken.⁶ Harsanyi argued that the distributive principle which is the rational choice for a self-interested person behind the veil of ignorance is the utilitarian principle that goods such as wealth should be distributed in such a way as to give each person the best chance of maximising their welfare. This conclusion, however, is not that which Rawls wanted to reach: his second principle of justice implies that wealth should be distributed in a way which is most beneficial to the least advantaged members of society, and Rawls argues that this principle is preferable to Harsanyi's utilitarian principle precisely because it affirms the value of reciprocity within society (*TJ* 102: 88). Yet Harsanyi's principle is that which is implied by Rawls's own original position argument because of the emphasis there on rational self-interest. Rawls rejects this conclusion on the grounds that the rational strategy to be adopted in the original position is the 'maximin' strategy of minimising the risk of faring badly; for, he argues, when a reflective person make his choice

of principles he has to allow for the possibility that he is designing ‘a society in which his enemy is to assign him his place’ (*TJ* 152: 133). We can certainly agree that if this possibility is to be assumed to be likely, then self-interest dictates that maximin would be the rational strategy. But Rawls explicitly denies that this eventuality is to be assumed to be likely; his claim is only that because this possibility has to be allowed for, the pursuit of rational self-interest behind the veil of ignorance favours a maximin strategy. The reply to this, however, is that given the unlimited general knowledge of the world which is assumed to be available in the original position (*TJ* 137-8: 119), the reflective person motivated by rational self-interest has no good reason to adopt a maximin strategy. He should aim to maximise his expected goods in the light of his knowledge of the world and thus opt for Harsanyi’s utilitarian principle.

In *A Theory of Justice* Rawls systematically uses his conception of the original position to present his theory of justice. If the argument above is correct this was a mistake. But it does not follow that the thesis which motivated the conception of the original position, namely that the requirements of justice are determined by principles which would be agreed by parties who recognise each other as free and equal, is also a mistake. There is, however, a separate issue here, namely how a hypothetical agreement can determine the actual requirements of justice. After all, Rawls himself remarks that ‘we cannot say that a particular state of affairs is just because it could have been reached by following a fair procedure A fair procedure translates its fairness to the outcome only when it is actually carried out’ (*TJ* 86: 75). Thus it looks as if Rawls’s principles can only have the status of principles of justice for the envisaged hypothetical situation; whereas the intended conclusion was that they are principles of justice for us here now. But how does justice in that possible world determine justice for us in the actual world?

The only way of dealing with this problem is, in effect, to affirm it and accept that justice, and morality in general, are modal concepts whose application in the actual world depends upon their application in possible worlds in which the requirements of reciprocity are fulfilled. Rawls does not put the point in quite this way, but he reaches what is essentially this conclusion in his 1980 Dewey lectures ‘Kantian Constructivism in Moral Theory’. Rawls’s constructivism is the view that because justice and morality in general capture the requirements of living within a society of rational persons who recognise each other as free

and equal, the requirements of justice for us, here and now, are ‘constructed’ in the sense that they are rationally determined by the implications of a hypothetical construction:

Apart from the procedure of constructing these principles, there are no reasons of justice. Put in another way, whether certain facts are to count as reasons of justice and what their relative force is to be can be ascertained only on the basis of the principles that result from the construction. (‘Kantian Constructivism in Moral Theory’ CP 351)

This position is ‘Kantian’ because it is inspired by Kant’s account of the role of practical reason in determining the moral law; but Rawls is careful to distance himself from the metaphysical aspects of Kant’s conception of practical reason – his philosophy is a naturalised Kantianism. ‘Kantian constructivism’ therefore joins ‘justice as reciprocity’ as the fundamental themes of Rawls’s moral and political philosophy.

Finally, I return to my starting point, Rawls’s conception of freedom. Rawls holds that freedom has three fundamental aspects.⁷ First, to be free is to be a responsible agent, someone capable of appreciating reasons for action and taking responsibility for what they do. For Rawls this capacity is grounded in ‘the moral power to have a conception of the good’, by which he means both that one has ‘a conception of the good’, that is, some fundamental values and attachments from which one gets one’s sense of identity, and that one has the ability to appraise and revise these values and attachments in the lights of reasons to do so. This moral power both defines the second aspect of our freedom and grounds its third aspect, which is that people are ‘self-originating sources of valid claims’. Reciprocity is then the mutual recognition of each other as ‘free and equal’ in this sense. As will be apparent, this freedom has nothing to do with any ‘natural’ freedom to withdraw from society: instead it is the freedom of someone whose capacity to construct a meaningful life of their own entitles them to full and proper consideration when moral and political questions are at stake. This is the ethical ideal which informs Rawls’s moral and political theory.

4. Basic liberties and the first principle of justice

In *A Theory of Justice* Rawls stated his first, fundamental, principle of justice as the principle that each person has a right to the most extensive scheme of basic liberties compatible with everyone else in their society having a right to similar liberties. In *Political Liberalism* Rawls modified this principle, substituting ‘a fully adequate scheme of basic liberties’ for the earlier talk of ‘the most extensive scheme’ of them (*PL* 291). What lies behind this change is

Rawls's acceptance of H. L. A. Hart's criticism of the principle as originally formulated. Hart had argued that Rawls's talk of the 'most extensive scheme of liberties' was unsatisfactory, since what is important as far as justice is concerned is the identification and protection of basic liberties, not their maximal extension, whatever that might be.⁸ Rawls agreed with Hart on this point and modified his first principle accordingly to affirm that justice requires the protection of 'a fully adequate scheme of basic liberties which is compatible with a similar scheme of liberties for all' (*PL* 291).

This revised principle raises the question which liberties belong to a 'fully adequate scheme', and this question cannot be separated from the question why this principle is the fundamental principle of justice in the first place. One of the odd features of *A Theory of Justice* is that Rawls does not address this question until he gets to the final chapter of the book, where he argues that liberty gets its priority from 'the central place of the primary good of self-respect' (*TJ* 543: 476⁹) because 'the basis for self-esteem in a just society is not then one's income share but the publicly affirmed distribution of fundamental rights and liberties' (*TJ* 544: 477). Although this claim is intelligible, it is not well integrated into the main structure of his theory. It is, therefore, no surprise that in *Political Liberalism*, as well as reformulating the principle, Rawls provides a much better rationale for its priority which is directly attached to the ideal of reciprocity. Since, as we have just seen, freedom includes 'the moral power to have a conception of the good', the recognition of each other's freedom will lead people to value the liberties which enable them to exercise this power. For Rawls there are two such basic liberties: liberty of conscience and freedom of association (*PL* 310-3). It seems to me that these liberties should be combined with rights such as the right to privacy and family life which are also integral to one's sense of one's own identity; but Rawls nowhere mentions this familiar line of thought. Instead he argues that the other basic liberties are political liberties which are to be justified by reference to another fundamental moral power, 'the sense of justice' (*PL* 302).

Rawls introduced the conception of the 'sense of justice' in *A Theory of Justice* in the course of a subtle account of development of moral sensibility. He argues that our sense of justice is the result of a different kind of reciprocity:

Because we recognize that <other persons> wish us well, we care for their well-being in return. Thus we acquire attachments to persons and institutions according to how we perceive our good to be affected by them. The basic idea is one of reciprocity, a tendency

to answer in kind. Now this tendency is a deep psychological fact. Without it our nature would be very difficult and fruitful social cooperation fragile if not impossible. (*TJ* 494-5: 433).

This form of reciprocity is not the reciprocity of ‘justice as reciprocity’, but they are closely related because the ‘tendency to answer in kind’ implies that where people find themselves treated as free and equal they are likely to treat others in the same way and in this way develop a sense of justice. Hence we are to think of the sense of justice as a normal disposition of anyone who has grown up in a reasonably just environment. For this reason, Rawls holds, it is a basic moral power, one which is central to our willingness to act towards others in ‘reasonable’ ways, just as our capacity for a conception of good is central to our self-interested ‘rationality’.

Rawls’s account of the development and role of a sense of justice is plausible and valuable. But it does not follow that a sense of justice is the ground for the status of political liberties as basic liberties which are essential constituents of a fully adequate scheme. For a sense of justice is essentially a disposition to fulfil the requirements of justice, whatever they are, and it does not of itself dictate what these requirements are to be even if it is dependent upon our ‘tendency to answer in kind’. Instead the status of these political liberties follows from an aspect of justice as reciprocity concerning which Rawls says surprisingly little in this context, namely equality. It is central to Rawls’s theory that justice is the working out of the recognition of each other as ‘free *and equal*’ (my italics); and the route from the recognition of each other as equal members of political society to some form of representative democracy which includes universal political liberties is familiar and well-established. Rawls, like Kant, is, I think, over-inclined to let his argument be dictated by the architectonic of his general position. He connects the political liberties with our ‘reasonable’ sense of justice in order to balance the connection between personal liberties and the ‘rational’ capacity for a conception of the good. But this schema overlooks the fact that the public, political, requirements of the just exercise of power within a society need to be determined separately from a willingness to do what justice requires. What is fundamentally wrong with Plato’s elitist constitution in the *Republic* is not that it violates the sense of justice of its citizens, but that it divides up humankind into groups with different ethical status. To recognize each other as free and equal is precisely to reject any such division and for this reason to demand equal political rights for all.

There is, however, a different kind of equality whose importance Rawls does acknowledge in this context, namely that the basic liberties identified by the first principle should be of 'equal worth' to all members of society. The main anxiety here arises from the effects of social and economic differences between people. In *A Theory of Justice* Rawls argued that this anxiety is adequately addressed by his second principle of justice which implies that these differences are only legitimate where they are of benefit to the least advantaged members of society. But he later accepted the need to provide a separate way of substantiating the equal worth of liberty: in *Justice as Fairness* (2001), his late restatement of his theory of justice, he remarks:

This principle <i.e. the first principle of justice> may be preceded by a lexically prior principle requiring that basic needs be met, at least insofar as their being met is a necessary condition for citizens to understand and to be able fruitfully to exercise the basic rights and liberties. (*JF* 44 note 7)

5. Social and economic inequalities and the second principle of justice

I turn now to the second principle of justice, that social and economic inequalities are to be arranged in such a way that positions of power and status are open to all and differences in wealth and income arise only through distributive systems which work if not to the advantage of everyone at least to the advantage of the least well-off members of society. This principle has two clauses, one concerning social inequalities, the other economic ones; Rawls says that the social clause takes priority over the economic clause, and I shall discuss them in this order.

The substance of the social clause is that justice requires that there be 'fair equality of opportunity', by which Rawls means that 'those with similar life abilities and skills should have similar life chances' (*TJ* 73: 63). The familiar method of realizing this requirement is through education, and Rawls holds that school systems should be organised so as to remove class barriers; in his later work he also maintains that basic health care needs to be assured for all, which addresses a further cause of unequal opportunities (*LP* 50). Rawls recognises, however, that one's life chances are also profoundly shaped by one's family; hence, he asks (*TJ* 511: 448), does justice require that the family be abolished? In *A Theory of Justice* he replies that once we take account of the economic clause of his second principle, the differences between different families should be sufficiently mitigated to avoid this

conclusion. This reply seems fair enough, though it is striking that he does not appeal to his first principle of justice to defend the family via a right to privacy, as would be normal today. Rawls returned briefly to the issue of the family in his late writings.¹⁰ Again, his concern is with equality of opportunity, but now with a different focus on the traditional roles of women as wife and mother insofar as these tend to make it impossible for them to enjoy the same opportunities for fulfilling and rewarding careers as men. For Rawls, this situation is inherently unjust; the only fair division of labour between men and women in the family is a voluntary one.

The economic clause of Rawls's second principle of justice, generally known as the 'difference principle', states that systematic differences in wealth and income are justified only where they work to the advantage of the most deprived section of society (*TJ* 75: 65). In applying this principle, Rawls starts from the belief that the economy of a just society will be organised as a free market, given the greater efficiency of a market economy as compared with a command economy and also its compatibility with the basic liberties of the first principle and the requirements of fair equality of opportunity (*TJ* 272: 240-1). In endorsing a market economy, however, Rawls emphasizes that socialist systems of collective or cooperative ownership are not ruled out; the choice between these and a system of private ownership, or some combination of them, should be a pragmatic matter (*TJ* 280: 248). Whatever system is adopted, however, Rawls maintains that 'income and wages will be just once a (workably) competitive price system is properly organised and embedded in a just basic system' (*TJ* 304: 268). Hence differences in wages will reflect 'features of jobs that are significant on either the demand or the supply side of the market, or both', such as 'Experience and training, natural ability and special know-how' (*TJ* 305: 269); but if the system is in other respects just, then these differences will also be just. Indeed, the situation is for Rawls a paradigmatic case of pure procedural justice (*TJ* 304: 267).

The anxiety to which this account gives rise is that it dissipates Rawls's initial egalitarian presumption by permitting those with scarce skills to bargain hard for disproportionately large salaries. Rawls will reply that within a just society people will not be motivated to press their claims in this way: for they should be assumed to be committed to realising the ideal of constructing a society in which any systematic income differential from which one benefits oneself is only a means of bringing greater benefits to others, especially the most deprived section of society (*JF* 64, 76-7). Furthermore, Rawls's conception of a just

distributive system includes provision for redistributive taxation which aims ‘gradually and continually to correct the distribution of wealth and to prevent concentrations of power detrimental to the fair value of political liberty and fair equality of opportunity’ (*TJ* 277: 245). Thus Rawls’s hope is that once these are points are taken into account, the resulting system is that which has the best chance of enabling the most deprived groups of society ‘to make intelligent and effective use of their freedoms and to lead reasonable and worthwhile lives’ (*LP* 114). One may not be altogether persuaded by this line of thought; but it would be a great mistake to read Rawls as an apologist for contemporary capitalist society. In his late writings he emphasises the difference between the kind of ‘property-owning democracy’ he favours and ‘welfare-state capitalism’, the kind of political economy characteristic of contemporary Britain and the USA (he later remarks that it was ‘a serious fault’ of *TJ* that it failed to emphasize this contrast; see *JF* 139 note 5). Welfare-state capitalism permits ‘very large inequalities in the ownership of real property ... so that control of the economy and much of political life rests in few hands’; hence although ‘welfare provisions may be quite generous and guarantee a decent social minimum covering the basic needs, a principle of reciprocity to regulate economic and social inequalities is not recognised’ (*JF* 138). By contrast, in a property-owning democracy, the aim is not simply to assist those who need help, but to disperse wealth and capital in order to ‘put all citizens in a position to manage their own affairs on a footing of a suitable degree of social and economic equality’ (*JF* 139). Rawls ends with the hope that in a system of this latter kind, the problem of social exclusion, or the ‘underclass’ as he calls it, will be addressed insofar as it can be:

Under these conditions we hope that an underclass will not exist; or, if there is a small such class, that it is the result of social conditions we do not know how to change, or perhaps cannot even identify or understand. When society faces this impasse, it has at least taken seriously the idea of itself as a fair system of co-operation between its citizens as free and equal. (*JF* 140)

Rawls’s difference principle has attracted more critical attention than all the rest of his work put together. I have already mentioned, and endorsed, Harsanyi’s criticism of Rawls’s attempt to derive the difference principle from his original position. But I take this to be primarily a criticism of Rawls’s conception of the original position; it does not settle the question whether one should favour the difference principle over the utilitarian approach favoured by Harsanyi. Rawls’s objection is of course that utilitarianism fails the test of

reciprocity since it does not protect the claims of the most deprived from being set aside in favour of greater benefits to other, better off, social groups (*TJ* 33: 29). But this argument is disputable: utilitarians will argue that given the diminishing marginal utility of wealth, the goal of maximising average welfare will direct a society to prioritise improvements in the situation of the most deprived. And when one considers the detail of Rawls's proposed policies for implementing his difference principle, as sketched above, it is hard to see anything incompatible with the implications of an enlightened utilitarianism. A more fundamental criticism of Rawls's difference principle came from his Harvard colleague Robert Nozick.¹¹ Nozick argued that the way in which this principle seeks to direct the distribution of wealth and income conflicts with the fact that wealth and income are personal property which it is not society's legitimate business to redistribute in order to achieve a particular pattern of distribution. Personal property is a personal 'entitlement' over which society has no legitimate claim without the consent of the property-holder. Rawls's reply is that personal entitlements arise only within a just system of property rights (*JF* 72-3) and that there is nothing improper about a system which uses taxation to constrain property holdings in order to achieve a just distribution of wealth and income (*JF* 51-2). The only way in which such interference would be illegitimate is if the property rights in question were to derive from a basic liberty to acquire property. Rawls, however, denies that there is any such basic liberty (*JF* 114).

Nozick's criticism of the very idea of distributive justice is therefore misguided. But Rawls himself is guilty of a comparable mistake in suggesting that his own difference principle follows directly from the fact that a person's 'native talents' are undeserved. Rawls argued that because one has no underived moral entitlement to one's native talents, it is appropriate for society to be organised in such a way that everyone benefits from their use, especially those who have not been favoured by the natural lottery which distributes them. These talents may therefore be viewed, he says, as 'a social asset to be used for the common advantage' (*TJ* 107: 92), and in this way, he holds, 'we are led to the difference principle' (*TJ* 102: 87). Rawls's use of the phrase 'social asset' here was unfortunate, since it suggested that there is social ownership of personal abilities; but that unintended meaning can be set to one side. The important question concerns the relationship between the difference principle and the fact that native talents are not deserved. The fact that native talents are undeserved certainly removes a Nozick-inspired objection to the difference principle. But the converse

implication does not hold: the fact that native talents are undeserved implies next to nothing about the just distribution of the goods people produce by making use of their talents. It is, for example, entirely consistent with a utilitarian distributive principle. Rawls's intuitive argument for the difference principle was a mistake; fortunately, nothing hangs on it.

6. The Rawlsian state

Rawls conceives of his two principles of justice as a characterisation of the 'basic structure of society'. So the Rawlsian state (if one can so speak) is to be thought of as essentially an implementation of these principles: the first principle frames the state's constitution and legal system, while second principle sets policy goals for the legislature. An important implication of the first principle is that the state is to be neutral as between the varied religious and ethical beliefs of its citizens except for those which conflict with the conception of justice as reciprocity which is fundamental to the state itself. Thus there can be no established church or other faith: as against defenders of confessional states, however tolerant, Rawls argues that it is unjust that those who do not share the official faith of their society should be disadvantaged with respect to the state's political institutions. When it comes to the implementation of the second principle of justice, however, the Rawlsian state is far from neutral. The first clause of the principle implies that the state should actively promote equality of opportunity in the fields of education, employment and so on; and the difference principle implies that the state must take an active role in promoting a property-owning democracy in which ownership of wealth and capital is dispersed.

What is not covered by the principles of justice is the duty of citizens to obey the laws of their state insofar as it is just. Since membership of one's society is not voluntary Rawls accepts that the fundamental ground for this duty cannot be a contractual obligation (*TJ* 335-6: 295-6). Instead, he says, it rests upon a 'natural duty of justice', understood as the duty 'to comply with and to do our share in just institutions where they exist and apply to us' (*TJ* 334: 293). For Rawls this natural duty is justified by the line of thought that is central to his conception of justice as reciprocity: in thinking about the rules which would be agreed by the members of a society whose members recognise each other as free and equal, he argues, we can see that these rules, whatever they are, would include the higher-order requirement to comply with whatever rules are agreed. (*TJ* 334-5: 295). It is, however, one thing to accept a duty to obey the law where the law is indeed just; but what where it is unjust? Rawls was

writing *A Theory of Justice* during the 1960's when the civil rights campaign led by Martin Luther King was at its height. It is not surprising, therefore, that Rawls added a discussion of civil disobedience to his account of political obligation, despite the fact that in doing so he was no longer dealing with the basic structure of an ideally just society, but was instead dealing with the requirements of justice in a situation in which there is serious injustice. Whereas traditional social contract theorists such as Locke had taken it that the appropriate response to extreme injustice ('tyranny') was rebellion and the removal by force of the existing state, Rawls argued that a political society which recognises the inescapable fallibility of its own political institutions should allow that there can be justified acts of disobedience which aim, not to overthrow the state, but to correct serious existing injustice:

Indeed, civil disobedience (and conscientious refusal as well) is one of the stabilising devices of a constitutional system, although by definition an illegal one. ... By resisting injustice within the limits of fidelity to law, it serves to inhibit departures from justice and to correct them when they occur. (*TJ* 383: 336).

Hence:

So understood a conception of civil disobedience is part of the theory of free government. (*TJ* 385: 338)

This was an important insight, previously propounded informally by Thoreau, Gandhi, and King, but here for the first time built into an account of the moral structure of a reasonably just state.

7. Political Liberalism

During the 1980's Rawls was brought to acknowledge the 'fact of reasonable pluralism', the fact that there are irresolvable disagreements concerning the foundations of morality between Kantians (such as himself), utilitarians, intuitionists, and adherents of the great variety of religious beliefs. Although he held that these disagreements do not entail scepticism, either about morality or about moral theory, they do imply that the public vindication of the requirements of justice which is needed to ground the stable allegiance of citizens to their society should not require the adoption of any one general ethical theory such as the Kantian theory he himself held. As a result Rawls was led to think that political philosophy requires the detachment of a 'political conception of justice' from any underlying ethical theory and his second major book, *Political Liberalism* (1993), was his attempt to present a conception

of this kind. This required a major change in his philosophical perspective from that of *A Theory of Justice*; but in many other respects the change is not great, since Rawls retained much of the substance of his theory.

In *Political Liberalism* Rawls still holds that the conception of justice is a ‘moral conception’ which applies to social institutions, but he now takes it that the moral requirements in question are the implications of values which are essentially political. And instead of relying on Kantian constructivism to develop these implications Rawls recasts his account of the justification of political judgments as ‘political constructivism’. This is the thesis that the requirements of justice can be justified by reference to ‘public reason’, the project of constructing publicly defensible rules for social cooperation among reasonable persons even if they have very different ethical beliefs. In this sense, ‘Political liberalism, then, aims for a political conception of justice as a freestanding view’ (*PL* 10). Despite these changes, however, Rawls’s general approach remained much the same, so that there was no inappropriateness in my introducing earlier his revised argument for the first principle of justice from *Political Liberalism*. But Rawls is now more concessive in his attitude to his principles:

Accepting the idea of public reason and its principle of legitimacy emphatically does not mean, then, accepting a particular liberal conception of justice down to the last details of the principles defining its content. We may differ about these principles and still agree in accepting a conception’s more general features. (*PL* 226)

By detaching his theory from comprehensive ethical theories Rawls aimed to be neutral with respect to such theories. Nonetheless he also holds that political constructivism can work only among the adherents of ‘reasonable’ ethical theories. But what does ‘reasonable’ mean here? One aspect is manifest in the doctrine of ‘reasonable pluralism’, where what is reasonable is what is defensible in the light of the evidence and thus meets the standards of theoretical reason (*PL* 56). But there is another aspect based on practical reason which concerns the willingness of people to be ‘reasonable’ with respect to one another. Here reasonableness is just reciprocity:

Reasonable persons, we say, are not moved by the general good as such but desire for its own sake a social world in which they, as free and equal, can cooperate with others on terms all can accept. They insist that reciprocity should hold within that world so that each benefits along with others. (*PL* 50)

In this sense, then, reasonable ethical theories are those which affirm the value of reciprocity. This conclusion is not surprising, but it shows how even when Rawls was aiming to advance a freestanding political conception of justice he retained his commitment to ‘justice as reciprocity’.

8 *The Law of Peoples*

Right from the start Rawls had envisaged that his theory of justice should be applicable to international affairs (JF 49); but in *A Theory of Justice* he provided only a very brief sketch of this application in which he claimed that it would just lead to the standard principles of the law of nations (TJ 378: 332). Towards the end of his life, however, he returned to this issue and his last major piece of writing, *The Law of Peoples* (1999), was an extended discussion of justice in international affairs. As the title of the book indicates, Rawls takes it that international affairs are best conceived as dealing with relationships between ‘peoples’, rather than nations, societies or states. This terminology is potentially misleading since whereas we normally speak of, say, the Kurds and Basques as peoples, he takes it that a ‘people’ will have a constitutional government with jurisdiction over an established territory, which the Kurds and Basques lack. It is in fact relations between states that he is concerned with, but for the purpose of exposition I will stick with Rawls’s term.

As one would expect, Rawls employs within international affairs much the same approach he had used in domestic affairs. The aim is to frame principles of justice, a ‘Law of Peoples’, to regulate relationships between members of a just international society, the ‘Society of Peoples’. These principles are to be thought of as founded upon the fundamental value of reciprocity:

Thus, the criterion of reciprocity applies to the Law of Peoples in the same way it does to the principles of justice for a constitutional regime. This reasonable sense of due respect, willingly accorded to other reasonable peoples, is an essential element of the idea of peoples who are satisfied with the status quo for the right reasons. (LP 35).

Not surprisingly, therefore, Rawls proposes that the Law of Peoples can be thought of as the rules which would be agreed to in a hypothetical original position in which the parties think of themselves as representing a people and coming to an agreement behind a veil of ignorance as to ‘the basic terms of cooperation among peoples who, as liberal peoples, see themselves as free and equal’ (LP 33). As ever, we do not have to endorse the details of this

to appreciate his general line of thought. Instead the focus needs to be on the proposed Society of Peoples and its Law.

One thing which is striking about Rawls's Society is that he allows that it includes 'decent' peoples which, despite not being wholly just liberal societies, respect basic human rights and the rule of law internally and refrain from aggressive war externally. It might at first appear that this represents a significant retreat from his defence of liberal politics. But Rawls argues that just as liberal political principles call for the toleration of non-liberal associations which pose no threat to others, a liberal Society of Peoples should tolerate decent peoples even if their constitution is in some respects illiberal (*LP* 59-60). This situation is, he argues, what we should expect of a 'realistic utopia'. The Society is a 'utopia' insofar as its basic principles represent the implications, rational and reasonable, of the basic value of reciprocity (*LP* 17-18); but it is also 'realistic' insofar as it acknowledges the fact of reasonable pluralism and the resulting diversity of decent peoples who subscribe at the level of international affairs to liberal principles (*LP* 18).

What then are the principles of international justice, the 'Law of Peoples'? As before Rawls largely reiterates existing principles of the law of nations (*LP* 37) and maintains that peoples are free and equal members of the Society of Peoples, with a right to self-defence, a duty not to intervene in the affairs of other peoples, and a duty to honour human rights internally. These points can be regarded as broadly equivalent to Rawls's first principle of (domestic) justice. But what is the international analogue of the second principle of justice which deals with social and economic inequalities? All that Rawls offers here is a duty of assistance:

Peoples have a duty to assist other peoples living under unfavourable conditions that prevent their having a just or decent political and social regime. (*LP* 37)

Rawls thinks of this duty as primarily owed to 'burdened peoples' which 'lack the political and cultural traditions, the human capital and know-how, and, often, the material and technological resources needed to be well-ordered' (*LP* 106). The aim, then, is to help such societies establish the conditions under which they can become well-ordered so that they can develop decent political institutions and a reasonably self-sufficient economy (*LP* 110-1). Rawls notes also the importance of ensuring that the terms of trade and other economic relations between peoples are fair (*LP* 42-3); but all this is very different in spirit from the difference principle. Rawls argues, however, that interventions to promote economic equality

between peoples are likely to be ineffective since the level of material prosperity of a people is primarily a matter of its own cultural traditions and values (*LP* 117). This claim is questionable. But Rawls has a further and better point, that it is essential to respect the independence of decent free peoples (*LP* 117). This point, however, shows that Rawls's approach to justice in international affairs was misconceived.

Rawls takes it that justice in international affairs 'simply extends' the ideas which inform his theory of justice for political societies (*LP* 123). Yet whereas he accepts that political societies should have a government with coercive powers (*PL* 136), he rejects the suggestion that the Society of Peoples needs a world government with coercive powers (*LP* 36). Rawls is right about this last point; but this fundamental difference between the two cases indicates that an understanding of international justice cannot be achieved by the simple extension to international affairs of a theory of justice for political societies.

Selected writings by John Rawls, with abbreviations:

‘Justice as Fairness’ (JF), *Philosophical Review* 67 (1958) 164-94. Reprinted in *CP* 47-72.

A Theory of Justice (TJ), Oxford: Oxford University Press, 1971. Revised edition 1999.

‘Justice as Reciprocity’ (JR), in *John Stuart Mill: Utilitarianism, with Critical Essays* ed. S. Gorowitz, Bobbs-Merrill: Indianapolis, 1971. Reprinted in *CP* 190-224.

‘Kantian Constructivism in Moral Theory’, *Journal of Philosophy* 77 (1980) 515-72.

Reprinted in *CP* 303-58.

‘Justice as Fairness: Political not Metaphysical’, *Philosophy and Public Affairs* 14 (1985) 223-52. Reprinted in *CP* 388-414.

Political Liberalism (PL), New York: Columbia University Press, 1993.

‘Fifty years after Hiroshima’, *Dissent* (1995) 323-27. Reprinted in *CP* 565-72.

‘The Idea of Public Reason Revisited’, *University of Chicago Law Review* 64 (1997) 765-807. Reprinted in *CP* 573-615.

The Law of Peoples (LP), Cambridge MA: Harvard University Press, 1999.

Collected Papers (CP), Cambridge MA: Harvard University Press, 1999.

Justice as Fairness: a Restatement (JF), Cambridge MA: Harvard University Press, 2001.

¹ 'Fifty Years after Hiroshima', in *CP*.

² The original, 1971, edition of *TJ* runs to 600 pages. In 1999 Rawls brought out a revised edition of the book for which the text was reset, with the result that the page numbers are different. I give page references in the form '(*TJ* x: y)', where 'x' denotes the page number in the 1971 edition and 'y' that in the 1999 edition. Some significant differences between the editions are mentioned in the notes.

³ I take this story from Brian Barry's obituary of Rawls in the *Financial Times* November 28, 2002.

⁴ The list of Rawls's writings at the end explains the abbreviations used here and elsewhere.

⁵ This passage is an addition in the revised edition.

⁶ J. Harsanyi 'Can the maximin principle serve as a basis for morality? A critique of John Rawls's theory' *American Political Science Review* 69 (1975) 594-606.

⁷ See 'Justice and Fairness: Political not Metaphysical' in *CP*, esp. pp. 404-8.

⁸ See H. L. A. Hart 'Rawls on Liberty and Its Priority' *University of Chicago Law Review* 40 (1973) 534-55; reprinted in *Reading Rawls* ed. N. Daniels (Blackwell, Oxford 1975) 230-52.

⁹ This and the next passage are altered in the revised edition of *TJ*; indeed in this edition the whole section, §82, is considerably altered.

¹⁰ See 'The idea of public reason revisited' in *CP*, esp. pp. 595-601. Rawls is here responding to Susan Okin *Justice, Gender and the Family* (Basic Books, New York, 1989).

¹¹ See *Anarchy, State and Utopia* (Blackwell, Oxford, 1974), esp. ch. 7.