

## A THEORY OF JUSTICE?\*

**ABSTRACT.** This is a critical analysis of John Rawls's *A Theory of Justice*. Rawls offers a theoretical justification of social democratic principles of justice. He argues that they are the principles which rational men would choose, under defined constraints, in an original position of social contract. The author criticises Rawls's assumption that men of any background, of any socialisation, would choose these principles in the original position. He argues that the choice which Rawls imputes to his contractors reflects a specific socialisation – one dominant in Western democracies. The theory is useful because it systematises a particular sense of justice; it is in no sense however a universal theory.

### I

My intention is to show that the contractual theory of justice defended by John Rawls does not have the status of a universal theory (John Rawls, *A Theory of Justice*, Oxford 1972 – henceforth 'J'). By a universal theory, I mean a theory which people in different circumstances, particularly people in different cultures, would have equal reason to accept – granted that they could all understand the argument for it. I intend to show that the most Rawls can claim is that his theory explicates the sense of justice of people in a particular society.

The paper has two subsidiary goals. The first is to suggest that the society for which Rawls provides a theory of justice is Western democracy, particularly in its twentieth century form – WD, for short. Rawls appeals to *our* intuition when WD nicely sums up what *we* have in common. Also he takes as natural attitudes which, if not exclusive to WD, are at least characteristic of it. The second subsidiary goal is to suggest that at the specific level of Rawls's argument only a particular theory of justice is possible; a universal theory would be something quite distinct.

Rawls is concerned mainly with the principles of social justice: "they provide a way of assigning rights and duties in the basic institutions of society and they define the appropriate distribution of the benefits and burdens of social cooperation" (J4). The basic institutions are "the political constitution and the principal economic and social arrangements" (J7). The particular principles for which Rawls argues are: first, "Each

person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all" (J250); second, "Social and economic inequalities are to be arranged so that they are both (a) to the greatest benefit of the least advantaged and (b) attached to offices and positions open to all under conditions of fair equality of opportunity" (J83; for a later reading see J302). These principles are interpreted under the constraints of two priority rules: roughly, that the first principle may never be compromised out of consideration for the second and that, in the case of the second, fair equality of opportunity may never be restricted out of consideration for the greatest benefit of the least advantaged (J302–03).

Rawls argues for these principles of justice by a version of the contractual theory. "They are the principles that free and rational persons concerned to further their own interests would accept in an initial position of equality as defining the fundamental terms of their association" (J11; see also J118–19). Rawls does not resort to a quasi-historical myth of a state of nature to give substance to his idea of the original position. It is a purely hypothetical situation defined by certain constraints and, taking account of the constraints, we are meant to be able to simulate the reflections of the imaginary contractors (J120).

The constraints which define the original position include constraints on the parties – they are to be rational, representative of possible social positions, mutually disinterested, reliable when it comes to complying with the principles and so on; these, and constraints on the task in hand – the principles to be chosen are principles of social justice, the society in which they are to apply is one of moderate scarcity, the principles are to satisfy such formal constraints as generality and publicity, etc. (J146–47). The central constraint however, is that of the veil of ignorance. This requires that the parties to the contract be in ignorance of their particular talents and fortunes in the society for which they are choosing principles, and indeed be in ignorance of the particular historical circumstances of that society – their knowledge extends only to general facts of politics, economics and psychology (J137).

The original position, by the present argument, is not a device which enables the theorist to step outside the limits of his place and time in history. It does not give him a voice to speak for men of cultures far removed from his own. This its inventor fails to appreciate: "to see our

place in society from the perspective of this position is to see it *sub specie aeternitatis*: it is to regard the human situation not only from all social but also from all temporal points of view” (J587).

## II

There are two lines which my argument might take. The first I will mention but not develop. It is the argument that the very idea of the original position makes sense only against the background of a certain social experience. It presupposes the experience of a society where the distribution of social and economic goods is regarded as something subject to human agency (1). More deeply perhaps, it presupposes the experience of social mobility and the uncertainty that this brings with it; otherwise the veil of ignorance constraint would seem quite outlandish.

The point I am making is reminiscent of C. B. MacPherson’s claim that the state of nature of which Locke (or Hobbes) spoke reflected a society in which market relations are dominant so that “the individual with which he starts has already been created in the image of market man” (*Possessive Individualism*, Oxford 1962, p. 269). I am arguing that the contractors with which Rawls starts also show signs of socially specific modelling: they are limit cases of socially mobile individuals considering a problem characteristic of a society that allows some economic intervention by the state. If this is so then some doubt is cast on the universal status of Rawlsian theory.

I do no more than mention this point because I do not think that it can be pressed home. The contractual theorist can say that though the original position is modelled on a situation specific to a certain society, there is no logical reason why it should not make sense in other societies. It is true that there is no reason in logic why it should not do so – but there is every psychological reason why it should not. The idea of the original position is too much the flower of one society – as I see it, WD – to have much chance of blooming in others, at least in some others. In the fourth section I shall mention a further consideration which suggests this view.

The line of argument which I prefer to follow now concentrates on the details of the choice attributed to the contractors. The choice is that of the two principles of justice and it is presented as “the unique solution to the problem set by the original position” (J119). I wish to argue that

Rawls makes a case for this conclusion which there is no reason to think would carry equal weight in all societies.

Rawls's book falls into three parts: on 'theory', 'institutions' and 'ends'. In each of these parts he provides a description which is meant to justify the choice he attributes to the original contractors. In the first part he describes the choice as the implementation of the maximin decision procedure, in the second as the adoption of principles in reflective equilibrium with our considered judgments of justice and in the third as the rational choice in view of the contractors' conception of goodness. The three descriptions recur in the book, but each is defended in its respective part.

### III

The maximin procedure is defined by a conservative rule for choice under uncertainty. "The maximin rule tells us to rank alternatives by their worst possible outcomes: we are to adopt the alternative the worst outcome of which is superior to the worst outcomes of the others" (J152–53). What Rawls wishes to argue is (a) that the choice of his two principles of justice by the parties in the original position is a maximin procedure and (b) that as such it is the appropriate procedure in the original position. He puts some faith in this argument: "if the original position has been described so that it is rational for the parties to adopt the conservative attitude expressed by this rule a conclusive argument can indeed be constructed for these principles" (J153).

About (a): Rawls can show that choice of his principles would be a maximin procedure only by comparing those principles with alternatives. He finds the alternatives in "a short list of traditional conceptions of justice" (J122). The conceptions he considers are mainly variants of utilitarianism: in particular he considers the principle that average utility – which, unlike aggregate utility, is insensitive to size of population – should be maximised. Here it is already clear that the reflections with which the contractors are credited are not very radical anthropologically: they are reflections traditional in WD. But perhaps the difficulty of the task does excuse these "rough and ready methods" (J123).

Rawls now has to show that choice of the two principles represents the maximin procedure, that the worst outcome of these two principles is better than the worst outcome of alternative procedures. His case has

some plausibility. A high average utility might mark a society which tolerated slavery and the worst outcome for an individual in such a society – being a slave – would seem to be worse than anything allowed in a society ordered by the two principles of justice; the criteria for this assessment of slavery certainly reflect a specific social experience, but I am willing to grant that they might be accepted by people of any background. The straight principle of average utility however is not the only utilitarian alternative to Rawls' principles. J. E. J. Altham has suggested this further alternative to me: "maximise average utility, subject to first having satisfied needs". It is not at all clear that the worst outcome of this would be worse than the worst outcome of the two principles of justice. I am not anxious however to get stuck on this point. Let us grant that choice of the two principles does represent the maximin procedure.

The second point which Rawls has to establish is that in the original position, the maximin procedure is indeed the rational one. He draws on economic decision theory and argues that in any situation there are three features which we must obtain if 'maximining' is to be the rational course. They are, that the situation be one of uncertainty within which probabilities cannot be assigned to the possible outcomes of any plan; that the minimum outcome promised by the maximin rule be satisfactory; and that some of the outcomes of alternative plans be unacceptable (J154). The Altham alternative does not involve any unacceptable outcomes but again I put this aside. Let us assume that these three conditions are realised in the original position.

This brings us to the central question. Granted that maximinning does mean choosing the two principles of justice, granted too that the original position satisfies the conditions which normally make maximinning the rational course, would choice of the two principles necessarily seem the rational choice for the contractors to make? How would it seem to someone of an aristocratic society for example, someone who found duelling regrettable perhaps but on many occasions the only reasonable course? None of us can say for sure. And that is precisely the point. It is only if we presuppose in the contractors a certain attitude to risk – one which contrasts the prudent and the imprudent, not the mean and the manly – that the maximin rule will seem the rational procedure for them to adopt. We readily make that presupposition because in WD we are generally disposed to take a conservative attitude to risk.

Rawls claims, contrary to this objection, that his argument does not rest directly on the assumption of an aversion to risk among the contractors. "What must be shown is that choosing as if one had such an aversion is rational given the unique features of that situation irrespective of any special attitudes to risk" (J172). This begs the question as it supposes that there is some independent test of the rational. At this stage, the only test is whether we can simulate the reflections of the parties in adopting the maximin procedure and feel their weight. This we can do with ease but only because of our characteristic attitude to risk. A little historical imagination enables us to see that others might find a very different rule of choice the natural one for the contractors to follow.

There is only one situation – one version of the original position – in which it is hard to imagine that people of any society would find the conservative rule of choice unnatural. This would arise if each contractor had to assume, not that he might, but that he definitely would be in the least advantaged position in the society to come. In this case the principle 'I cut, you choose' would apply (2). It would leave the choice of the maximin alternative contingent, not on a particular attitude to risk, but on an attitude to coming out worst. This would be an advantage insofar as aversion to coming out worst seems to be a more natural and universal human attitude than aversion to risk-taking. It would undermine the idea of a contract however: a contractor could not assume that he would be in the least advantaged position if he recognised that every other contractor made the same assumption.

#### IV

The second description which Rawls gives of the choice which he imputes to the contractors presents the principles chosen as principles in reflective equilibrium with our considered judgments of justice. They are in equilibrium because the conclusions which they yield about how to organise the basic institutions of a society are in line with our intuitive judgments; the equilibrium is described as reflective because it gives us a view of the grounds and mode of derivation of those judgments (J20).

Rawls does not say that the principles must in every case generate conclusions in line with our judgments. In central cases they will. In marginal cases where we are not sure of our judgments they may not: here the conclusions may shift our intuitive judgments somewhat, providing us with

independent theoretical guidance (J19–20). The theory of justice is Socratic: “we may want to change our present considered judgments once their regulative principles are brought to light” (J49). The important point however is that in the main, the theory fits our considered judgments. This fit is invoked by Rawls to justify the two principles which he takes the contractors to choose. “One can work out their consequences for institutions and note their implications for fundamental social policy. In this way they are tested by a comparison with our considered judgments of justice” (J152).

I do not intend to go into the consequences which Rawls draws from his principles: they are consequences which describe WD in fair outline. What I want to point out is that the test of reflective equilibrium is inconsistent with a ‘universalist’ view of the theory of justice. This test ensures that what Rawls’ theory does is explicate the sense of justice dominant among those whose judgments are taken into account, systematise the principles that they find natural and attractive. There is no more reason to believe that one theory of justice would satisfy all cultures than there is to believe that all cultures share the same judgments of justice. Rawls provides a theory which answers only to our judgments, a theory of justice for WD. (3)

Rawls does try to give scientific status to the test of reflective equilibrium, and this may seem to do something for the status of his theory of justice. He points out – rightly, I think – that such a test also appears in linguistics, at least as described by Noam Chomsky. “In this case the aim is to characterise the ability to recognise well-formed sentences by formulating clearly expressed principles which make the same discriminations as the native speaker” (J47). Here too the theory is Socratic, the principles may influence our intuitive judgments: “while we may not expect a substantial revision of our sense of correct grammar in view of a linguistic theory the principles of which seem especially natural to us, such a change is not inconceivable, and no doubt our sense of grammaticality may be affected to some degree anyway by this knowledge” (J49).

The analogy with linguistics however is a give-away. The analogy Rawls draws is with a particular grammar, the grammar of a particular language. It suggests that what he offers himself is a particular theory of justice, the theory which explicates a particular sense of justice. He might have found a less embarrassing analogue in universal grammar, the theory of

the constraints under which all particular grammars must work. The trouble is that his theory does not have the required similarities with that grammar. It just is not a universal theory of justice.

A further interesting point is suggested by the linguistic analogy: it links up with the argument mentioned in the second section but not developed. In Chomsky's linguistics, the descriptive adequacy of a grammar is its capacity to generate all and only the grammatical sentences of a language – the fact that it is in reflective equilibrium with our sense of grammaticalness. The drawback with descriptive adequacy as a criterion of a grammar however is that there may be two grammars which are capable of generating the grammatical sentences of a language. Chomsky finds a further criterion, a procedure for evaluating two descriptively adequate grammars, in universal grammar: the one which fits in better with the universal theory of language is the better grammar. It now has explanatory adequacy as well as descriptive (*Aspects of the Theory of Syntax*, Cambridge, Mass. 1965, pp. 26–27). The question is, does Rawls have any corresponding procedure for choosing between two theories of justice?

He needs such a procedure, for it is at least conceivable that another theory of justice should also meet the test of reflective equilibrium. Why should a theory not be possible which set up a divine legislator and gave him such attributes that his imagined judgments on justice would be in reflective equilibrium with our own? How would Rawls decide then between such a theory and that which he defends himself? Clearly he would decide for his own; so, I suspect, would all of us. But why? The only reason I can see, in the absence of a universal theory of justice which would do the work of Chomsky's universal theory, is that Rawls' theory is a native flower in WD, the alternative an exotic growth; here I return to the argument of the second section. A myth of rational contractors in an original position is a more plausible myth to our minds than the myth of a divine legislator.

v

The third description which Rawls gives of the choice of the two principles of justice presents it as the rational choice for the contractors in view of their conception of goodness. Does this description do any more than the others to give universal status to the Rawlsian theory of justice? I am



going to argue that it does not because I do not think that Rawls establishes the validity of this description.

First, what does he mean by 'rational'? He takes his cue from the economic theory of decision-making. "Thus, in the usual way, a rational person is thought to have a coherent set of preferences between the options open to him. He ranks these options according to how well they further his purposes; he follows the plan which will satisfy more of his desires rather than less, and which has the greater chance of being successfully executed" (J143). Here Rawls gives us the three main principles that he thinks of as guiding rational choice: the principles of effective means, inclusiveness and greater likelihood (J411–12). The principles are useful because they enable us to compare different plans; they may even be allowed a universal appeal and validity. What they do not do however is mark out one plan as the rational one in a situation of choice – at least not necessarily. Hence they cannot in themselves determine the choice of the original contractors.

What more is necessary in a situation of choice if one plan is to stand out as the rational plan? First, Rawls says, that the requirement of deliberative rationality is met. A plan is chosen with deliberative rationality when all its consequences have been clearly foreseen and balanced against the consequence of alternative plans (J417). Even when this requirement is met, however, it is not possible to say what is the rational plan for an individual in a situation of choice; that plan has only been specified formally. What it is also necessary to know is the desires which weigh with that individual. Thus if we are to be convinced that the rational plan for the original contractors to choose is represented by the two principles of justice we must be told something about their desires, their conception of the good.

The distinctive thing however about the choice which faces the contractors is that they do not, and in simulating their reflections we do not, know anything of their particular desires; this, by the veil of ignorance condition. The contractors are to choose principles of social organisation, not out of concern for such individually variable desires, but out of concern for desires to which they are bound, regardless of who they are. These are described as "general desires" (J263) or desires for "primary goods" (J93): they are meant to be desires for conditions required for the pursuit of any particular goals.

Rawls maintains, that these conditions, the primary goods, are certain rights and liberties, opportunities and powers, income and wealth and – something supposed to be made possible in the pursuit of these – self-respect (J92). He also maintains that these are ordered so that liberty is prior to socio-economic advantage and equality of opportunity to economic welfare (J302–03); other motivational assumptions, e.g. that incentive requires inequality, I leave unexamined. If his claims are allowed, then it certainly follows that Rawls’s third description is a fair account of the choice which he imputes to the contractors; this is indeed the rational choice for them to make in view of their desires, their conception of the good.

Does Rawls justify his claims, does he establish his list of the primary goods? He certainly sets out a strategy for doing so. They are meant to be the goods to which the contractors are directed in view of certain “general facts” about human beings (J424). These include facts about human desires and wants, capacities and abilities, and social interdependence: they are not discussed – on the grounds that they are matters of “common sense knowledge” (J425). The only fact discussed is described as a “deep psychological fact” (J432). “It says only that we prefer, other things equal, activities that depend upon a larger repertoire of realized capacities and that are more complex” (J429–30). Rawls calls it the Aristotelian principle.

It is worth noting one reason why Rawls may feel free to pass so quietly over facts that play a crucial role in his theory. It is that he shows a not unexpected faith in psychology, an optimism about the invariability of the facts which this discipline may be presumed to catalogue. Thus he speaks of general psychological principles of which he allows his contractors to be aware (J24, 456), principles which include “laws of motivation” (J26). In doing so he backs up his assumption that men of any society would attribute to the contractors in the original position the same psychology that he gives them. His positivism on this issue begs the question of whether the original position makes possible a theory to explicate a universal sense of justice.

But even if we grant Rawls his general facts – and what they are is not at all clear – we find that he fails to clinch his argument that they would lead the contractors to recognise the primary goods that he lists. At the point where we expect to have the argument clinched he falls back on

self-evidence. “We must assume, then, that the list of primary goods can be accounted for by the conception of goodness as rationality in connection with the general facts about human wants and abilities, their characteristic phases and requirements of nurture, the Aristotelian principle, and the necessities of social interdependence.... I shall not argue the case for the list of primary goods here, since their claims seem evident enough” (J434). Rawls’s failure to argue this case means his failure to establish the validity of his third description of the choice attributed to the contractors. It means that that description does no more than the others to establish the universal status of his theory of justice.

## VI

But in fairness: Rawls does offer an argument for the necessity of one primary good – that of self-respect. He defines self-respect as (a) having a sense of the value of one’s plan in life and (b) having confidence in one’s ability to carry it out. He appeals to intuition in defence of the idea that this is a primary good. “Without it nothing may seem worth doing, or if some things have value for us, we lack the will to strive for them” (J440). One can see the force of this appeal in some cases, that of a professional man in *WD* for instance, but one asks whether it applies in all. Rawls thinks that it does and argues that each person must find something he is good at – and, to satisfy the Aristotelian principle, it had better be of some complexity – and for which he earns the respect of some group of people, respect which encourages him in self-respect. “Thus what is necessary is that there should be for each person at least one community of shared interests to which he belongs and where he finds his endeavours confirmed by his associates” (J442; see also J178–79).

This definition of self-respect is tied in one obvious respect to a non-universal pattern of social experience, a pattern characteristically exemplified in *WD*. It is individualist. I suggest that for societies in which the category of the individual received less emphasis, for societies where social facts were not presumed in the norm to be the outcome of free individual actions, this definition would not do. The definition does not make it a condition of self-respect that one should see oneself as part of a larger entity – the group – in which one believes. This might well appear to be a condition of self-respect to somebody socialised in a more tightly

integrated society than WD. But Rawls leaves no room for such a person in his original position. His very definition of society – “a cooperative venture for mutual advantage” (J84) – must alienate anyone of a collectivist mentality from the thought experiment of the original position. To get contractors into the original position in the hope of simulating their reflections, it appears that we are to bring them through the purgatory of WD.

Rawls’ concept of self-respect is important because he tries to derive an argument from it for the priority – and indeed the basic appeal – of liberty. ‘Liberty’ includes a variety of freedoms: the freedom to vote and stand for public office, the freedom of thought, speech and assembly, the freedom of person and (personal) property, freedom from arbitrary arrest (J61); these, and also apparently “the important liberty of free choice of occupation” (J274). Rawls does not comment on the heterogeneous appearance of the list. In general, liberty is something generated by institutional arrangements – a “pattern of social forms” (J63) – and it has the merit of promoting self-respect.

Rawls argues that in the society which the contractors foresee equal liberty is a sure basis for self-respect, one which nobody would want to compromise, even in the hope of material advantage. It is a basis of self-respect because of what it allows: “the full and diverse internal life of the many free communities of interests” (J554; see also J442). And also, apparently, by a shift in the concept of self-respect, it is meant to provide such a basis in its own right. “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. And this distribution being equal, everyone has a similar and secure status when they meet to conduct the common affairs of the wider society” (J544).

Against this it must be said that the connection between liberty and self-respect cannot be put up as one that must command universal assent, even when self-respect is defined in an individualistic way. There can be many communities of interest, enough to give everybody a chance of self-expression, in a society which we would not regard as free. And there may also be sufficient matter for self-respect at the general political level. One might argue that in a society terrorised by a dictatorial power, one would not have much room for self-respect of any kind. But the argument can hardly go to show the need for full political liberty in the Rawlsian sense –

certainly not something like liberty of occupation. Liberty from arbitrary arrest may be put in the same list as other 'liberties', but this does not make it of a kind with them.

Rawls himself mentions the possibility of a feudal or caste society in which individuals find a source of self-respect in the role given them in the order of things. He says that this is an acceptable guarantee of self-respect only because of the assumption that the order of things is not subject to human choice, an assumption which is ruled out by the nature of the enterprise on hand (J547). On the contrary: the only assumption necessary is that the order of things in question has the most to be said for it in view of the primary goods recognised by the contractors. Only by begging the question can equal liberty be taken to be one of those goods.

## VII

The conclusion I draw is that Rawls does not establish that universal status which he assumes belongs to his theory of justice. The contractual device gives the principles of justice as output only because the input is a socially specific mentality – the mentality, I suggest, of people in *WD*. It will not do to say that the device gives even the general outline of a “finally adequate theory” (J581). It can only give the outline of a theory adequate to a particular sense of justice.

The reason for this should now be clear. It is not any dialectical incapacity in Rawls but a reason in principle. If a theory of justice is meant to generate judgments of justice satisfying the test of reflective equilibrium then it is going to be relative to the set of judgments, the particular sense of justice, with which it tries to achieve equilibrium. A theory of this kind cannot assume universal status. But there is no scandal here. I do not see what is wrong with saying that the sense of justice by which we order – or at least criticise – our society reflects the image of man we have created and not an eternal human essence. I do not see why theory should give itself the task of rescuing that sense of justice from history. The theory of justice is the means by which we explicate and examine our sense of justice, it is not a means of providing it with metaphysical foundations.

The linguistic analogy suggests that we should have a universal theory of justice as well as a particular; this would describe the constraints which any theory of justice must meet. The best candidate for the title of uni-

versal theory is the traditional philosophy of justice – the metatheory of morals and politics. What prescriptivists and descriptivists do for instance can be described in this way: they formulate conditions – of form and content – which every sense of justice – and, more generally, every moral sense – must meet, conditions which any particular theory of justice must respect.

In recent work on justice and morality, it is fair to say that we have had enough of universal theory. That is why Rawls's work is so interesting and so challenging; it offers a particular theory of our sense of justice. It is only a pity that Rawls himself should have confused it with the other sort of theory and tried to credit it with universal status.

*Trinity Hall, Cambridge*

#### NOTES

\* My warm thanks go to Jimmy Altham, Harry Bracken, Stefan Collini, Martin Hollis, Steven Lukes, Quentin Skinner and Denys Turner. Without their comments this paper would certainly have been a poorer effort. I am also grateful to Professor Rawls for a useful correspondence about his position.

<sup>1</sup> I owe this point to Jimmy Altham.

<sup>2</sup> I owe this point to Elizabeth Anscombe.

<sup>3</sup> In WD there is wide disagreement on particular judgments of justice between politically defined groups. I take it, with Rawls, that there is equally wide agreement on more general judgments, when the judgments are considered out of a political context.