



Incentives and Principles for Individuals in Rawls's Theory of Justice¹

By/Par

Alex Voorhoeve

Department of Philosophy, Logic, and Scientific Method
London School of Economics

ABSTRACT

Philippe van Parijs (2003) has argued that an egalitarian ethos cannot be part of a post-*Political Liberalism* Rawlsian view of justice, because the demands of political justice are confined to principles for institutions of the basic structure alone. This paper argues, by contrast, that certain principles for individual conduct—including a principle requiring relatively advantaged individuals to sometimes make their economic choices with the aim of maximising the prospects of the least advantaged—are an integral part of a Rawlsian political conception of justice. It concludes that incentive payments will have a clearly limited role in a Rawlsian theory of justice.

¹ I thank Peter Dietsch, Michael Otsuka, and Martin Wilkinson for comments on earlier versions of this paper.

INTRODUCTION

As is well known, John Rawls's restriction of the scope of the difference principle to what he calls "the basic structure of society"² plays a crucial role in justifying incentive-based income inequalities in his theory of justice. But what justifies this restriction of the scope of the difference principle? Philippe van Parijs (2003, 226-231) has argued as follows that this restriction reflects the demands of political liberalism.

Van Parijs believes that in *A Theory of Justice*, Rawls presents parts of a "complete conception of right", which is a conjunction of "justice as fairness" applying to the institutions of the basic structure (to which most of the work is devoted) and "rightness as fairness" applying to individuals' actions within the space permitted by these institutions (which is dealt with more briefly). It would be natural to think, Van Parijs remarks, that this complete conception of right must contain a principle requiring that individuals adopt an egalitarian ethos, given the potential of such an ethos to improve the situation of the worst-off and decrease inequality. The ethos will require that, within the constraints of a reasonable personal prerogative, relatively advantaged individuals make their economic choices with the intention of maximising the prospects of the least advantaged. (Note that this principle only becomes plausible once one expands the Rawlsian measure of advantage from primary goods to a measure of advantage which includes such goods as leisure, the burdens of work, and job satisfaction. Otherwise, those who are advantaged in terms of income might be required by the ethos to work very long hours, or perform work that they hate, in order to increase the income of those who work very little, or who hold low-paid, but very enjoyable jobs (see Cohen 2005). I will take it as given from now on that we are dealing with this expanded conception of advantage.) However, in *Political Liberalism* (xlii) and *Justice as Fairness* (xvii), Rawls argues that justice as fairness should not be seen as uniquely grounded in a particular comprehensive conception of right. Van Parijs believes that this implies that the Rawls of *Political Liberalism* must confine the principles of justice to principles for the institutions of the basic structure. For, he argues, requiring certain principles for individual conduct beyond those needed to support and maintain just institutions would be to cross the line from a political conception of justice—one that is independent of and without grounding in premises peculiar to metaphysical, epistemological, and general moral conceptions, and which can be shared by people who hold very different conceptions of these kinds—into a comprehensive moral conception.

² The basic structure is "a public system of rules defining a scheme of activities that leads men to act together so as to produce a greater sum of benefits and [that] assigns to each certain recognised claims to a share in the proceeds" (TJ, 74). Rawls's discussion suggests that these rules are formal, in the sense that they are spoken, written or tacitly understood statements that carry firstly, a known range of sanctions, second, a norm or rule that prescribes these sanctions, and thirdly, provisions for monitoring, all of which emanate from a rule-making arena like a court or government (see Frank Vandebroucke 2001, chapter VII). This understanding of the basic structure is consistent with Van Parijs's interpretation of it. It also makes for a relatively clear distinction between rules that are part of the basic structure and rules governing individual behaviour within the bounds permitted by this structure, since these latter rules need not be formal in the aforementioned sense.

My main purpose in this paper is to argue that Van Parijs's justification of the limited scope of the difference principle fails, because his idea that the 'political liberalism/comprehensive moral doctrine' divide is marked by the 'principles for the basic structure/principles for individual conduct' divide is mistaken. I do so in section 1 by showing that certain principles for individual conduct that go beyond individuals' duties to establish and support just institutions are an integral part of Rawls's political conception of justice because of their role in creating a stable society. In section 2, I explain why I believe that this review of the role of principles for individuals in Rawls's theory helps establish that this theory should be supplemented by an egalitarian ethos.

1.

Throughout his work, Rawls refers to the basic structure as the "primary" subject of social justice, but never writes that it is the sole subject of social justice (TJ, 6 and 47; PL, 11-12; JF, 10). Indeed, Rawls writes that a suitable specification of the principles that apply to the separate and free transactions between individuals and associations is essential to a theory of justice (TJ, 93ff and 293ff; JF 53-54). Four sections of *A Theory of Justice* (18, 19, 51 and 52) are devoted to discussing principles for individuals. The difficulty is, however, that Rawls discusses these principles in the context of a complete conception of right, which he calls "rightness as fairness", which he later describes as outside the bounds of a purely political conception of justice (TJ, 15 and 95-96; PL, xlii; JF, xvii). Does this mean that these principles for individuals have no place in Rawls's purely political conception of justice, as Van Parijs believes?

I do not think so. For the principles for individuals Rawls discusses in *A Theory of Justice* at any length all belong to the domain of political justice.³ Rawls considers certain principles for individuals part of social justice because these principles play an essential role in ensuring the stability of a just society. As Rawls writes when he introduces his lengthier discussion of principles for individuals:

"I now wish to take up the principles of natural duty and obligation that apply to individuals. The first two sections examine the reasons why these principles would be chosen in the original position and their role in making social co-operation stable" (TJ 293).

A society ordered by the principles of justice must be stable—must reproduce its conditions of existence, both materially and in its institutions and the attitudes and behaviour of its citizens—because it would be irrational for the contracting parties to agree to principles that

³ After mentioning the principle of fairness and the natural duty to support just institutions, Rawls writes: "I shall say very little about the other kind of principles for individuals. For (...) I must limit myself to the theory of social justice" (TJ, 100). See also JF (xvii): "(...) the problems examined in *Theory* in any detail are always the traditional and familiar ones of political and social justice".

could not be adhered to, since these would not be able to play the role justice is designed to play, i.e. ensure effective and fair co-operation in society (TJ, 125-126). Therefore, if they contribute to the establishment of social justice, the principles for individuals developed in *A Theory of Justice* must survive the shift to a purely political conception of justice in Rawls's later work.

I will look at three of the principles for individuals which Rawls discusses, and trace their relation to the problem of stability: the duty of mutual respect, the duty of mutual aid, and the principle of fairness.

The duty of mutual respect demands that individuals recognise and treat one another as beings with a sense of justice and a conception of the good (TJ, 155-156; 297). It implies a willingness to see things from another's point of view, and to give reasons to others whenever their interests are materially affected, thereby implying that citizens should form a community of justification. It disallows contempt or lack of esteem for others, and asks that citizens treat each other courteously, and are prepared to do each other small favours. The duty of mutual aid requires that we are to help others who are in need when doing so is not too costly for us (TJ, 100; 297-298).

Now, these two duties are not just important because of the balance of good over bad that would result from their being generally followed. Rather, they are part of the principles of social justice because of the way they affect the relations between individuals and individuals' self-respect. As Rawls writes:

“Once we try to picture the life of a society in which no one had the slightest desire to act on these duties, we see that it would express an indifference, if not disdain for human beings that would make a sense of our own worth impossible” (TJ, 298).

It is this connection with self-respect that makes these principles part of a theory of social justice. Rawls writes that parties in the original position would want to avoid at almost any cost the conditions that undermine self-respect (TJ, 386). One reason for this importance is that without a secure sense of self-worth, people's lives become meaningless to them: they can see no point in planning and managing their lives, or in undertaking any activity. A second reason is the role of people's sense of self-worth in establishing effective social co-operation, and it is this reason that concerns us here. Self-respect affects stability in two ways (TJ, 155). The first is that a society in which individuals support each other's sense of self-worth is a prerequisite for the development of citizens' sense of justice. Rawls's view of how citizens acquire an effective sense of justice depends on a three-stage theory of moral development (TJ, 429-430). At the first stage, children raised in a caring environment form affective ties with the members of their family. At the second stage, finding that institutions are just and that people generally and with evident intention act on principles of justice, people develop trust in these institutions and their fellow citizens. At the third stage, realising that their interests and the interests of the people they care about are furthered by

the arrangements of a just society, and their worth is affirmed, people acquire a normally effective sense of justice.

Now, the duties we are discussing impact either directly, or indirectly through their influence on the sense of self-worth, at all three stages of this development of the sense of justice. Firstly, Rawls argues, people with a secure sense of self-worth are more likely to care for their children (TJ, 436). Secondly, other people's willingness to take our needs into account and treat us with respect contributes to the process of coming to trust them (the second stage) and to the fact that social arrangements are such that they further our good and affirm our worth (the third stage).

The second way that self-respect influences the stability of society is through its relation to envy. Envy in the sense in which it involves ill-will towards the person envied is what Rawls calls a "disruptive attitude" because it is collectively disadvantageous (TJ, 125). It leads the worse-off to wish to deprive the better-off of their greater benefits, even at some cost to themselves. In turn, this leads the better-off to take precautions against the hostile acts to which the worse-off become prone. Conscious of others' negative attitudes towards their good fortune, they may even become spiteful: they become willing to give of goods of their own in order to deny the less well off certain benefits. People moved by envy and spite, then, will not have the motivation to adhere to fair and mutually beneficial arrangements. It is essential, then, that in a well-ordered society the conditions for strong feelings of the destructive form of envy do not arise.

Now, Rawls identifies lack of self-worth as one of the conditions under which envy may arise.⁴ And we have established that general adherence to the duties of mutual aid and mutual respect contributes to citizens' sense of self-worth. It follows that general adherence to these duties will tend to lessen the tendency for envy to develop in society and hence contribute to its stability.

The duties of mutual aid and respect are not restricted to people's actions within the institutions of the basic structure or their activities that pertain directly to the establishment and maintenance of a just basic structure. For one of their functions, to sustain mutual trust and confidence in one another's intentions, cannot be fulfilled if individuals apply them so selectively. Not to have or act on racist attitudes, for example, follows from the duty of mutual respect (PL, 195). It would seem both very strange and in contradiction with Rawls's stated purpose for taking this to be a duty to think that this requirement applies only when people are fulfilling their roles in one of the institutions that make up the basic structure of society, since racist incidents, in whatever context they occur, would clearly affect the quality of individuals' relations with each other. Rather, it seems natural for it to hold in people's everyday lives and therefore to conclude that it extends beyond the basic structure.

⁴ "[s]omeone sure of the worth of his plan of life and his ability to carry it out is not given to rancor nor is he jealous of his good fortune. Even if he could, he has no desire to level down the advantages of others at some expense to himself. This hypothesis implies that the least favored tend to be more envious of the better situation of the more favored the less secure their self-respect" (TJ, 469).

Let us now look at the principle from which Rawls believes all obligations derive, which he calls “the principle of fairness”. This principle holds that a person is required to do her part as defined by the rules of a legitimate institution if she has voluntarily made use of the institution to further her interests (TJ, 96). This principle may seem to ask for nothing beyond the behaviour necessary to maintain the basic structure. However, one principle that derives from the principle of fairness is not so restricted. What Rawls calls the “principle of fidelity” holds that promises that are made under the right conditions should be kept (TJ, 303-306). This principle applies to all voluntary relationships that people enter into. It is part of a theory of social justice because of its role in enabling and stabilising co-operative arrangements for mutual advantage and in building trust among citizens. This latter consequence ties it to the development of the sense of justice, in ways discussed above for the natural duties.

In sum, the duties of respect and mutual aid, and the principle of fidelity cannot be reduced to part of the basic structure. In addition, they are essential to establishing social justice as Rawls conceives it. It follows, contra Van Parijs (2003), that Rawls's principles of political justice are not restricted to the basic structure.

2.

This conclusion implies that, absent other arguments for limiting the scope of the difference principle, the reasons Van Parijs mentions for favouring an egalitarian ethos—that its adoption will improve the situation of the worst-off and decrease inequality—retain their force. I also believe our review of the role of the principles for individuals that Rawls does accept reveals another reason for favouring this ethos: its contribution to making social co-operation stable. This contribution runs through two routes. First, by diminishing inequality, it directly lessens the tendency for destructive envy to arise in society. Second, everyone's evident willingness to freely—that is, without coercion of the state—act to improve the prospects of the least advantaged is an affirmation of each individuals' worth, and especially of the worth of those individuals whose sense of self-worth might be undermined by being among the least advantaged in society. Thus, an egalitarian ethos would contribute to the stability of social co-operation through the same routes as the natural duties.

As is well-known, others have advanced different arguments for limiting the scope of the difference principle, based on the demands of publicity (see Andrew Williams 1998) and of freedom of occupation (see Brian Barry 1989). I believe, however, that these arguments have been successfully responded to: Vandenbroucke (2001) argues convincingly that an egalitarian ethos can be public, and Cohen (2005) argues forcefully that enjoying freedom of occupational choice in the sense of not being coerced by the government or by public pressure to choose a particular job is perfectly consistent with a moral duty to use one's freedom of occupational choice in ways that improve the situation of the most disadvantaged.

If this is correct, then a Rawlsian has, I believe, good reasons to endorse an egalitarian ethos. As a consequence, from the perspective of a Rawlsian theory of justice, we can answer the

question set for these essays—which income inequalities, if any, can be justified as incentive payments?—as follows. Income inequalities due to incentive payments are permissible if they (a) compensate for the burdens of labour (including things like training, health risks, long working hours, etc.) or for lesser job satisfaction and are consistent with achieving equality at the highest possible level, or (b) serve to induce individuals exercising their legitimate personal prerogative to make choices that are to the benefit of the least advantaged. Incentive payments of type (a) will not generate inequalities in advantage, whereas incentive payments of type (b) may.

BIBLIOGRAPHY

Works by John Rawls

- CP *Collected Papers* edited by S. Freeman. Cambridge: Harvard University Press.
JF *Justice as Fairness: A Restatement* edited by E. Kelly. Cambridge: Harvard University Press.
PL *Political Liberalism* (2nd edition). New York: Columbia University Press.
TJ *A Theory of Justice* (revised edition). Oxford: Oxford University Press.

Other references

- Barry, B. (1989) *Theories of Justice*. London: Harvester-Wheatsheaf.
- Cohen, G.A. (2005) *The Freedom of, and in, Occupational Choice Trilemmas*. Manuscript.
- Van Parijs, P. (2003) 'Difference Principles', in *The Cambridge Companion to Rawls*. Edited by S. Freeman. Cambridge: Cambridge University Press, pp. 200-240.
- Vandenbroucke, F. (2001) *Social Justice and Individual Ethics in an Open Society*. Berlin: Springer Verlag.
- Williams, A. (1998) 'Incentives, Inequality and Publicity', in *Philosophy & Public Affairs* 27, pp. 225-247.