B.J.Pol.S. **34,** 251–268 *Copyright © 2004 Cambridge University Press*
DOI: 10.1017/S0007123404000043 *Printed in the United Kingdom*

Just Constraints

ANDREW MASON*

Political theorists disagree about the extent to which issues of feasibility, stability, institutional design and human nature can be bracketed in analysing the concept of justice. At one end of the spectrum some argue that no analysis of justice can be adequate in the absence of an account of how it could be implemented, whereas at the other end there are those who argue that principles of justice are logically independent of issues of feasibility. Influenced by the work of John Rawls, many theorists occupy the middle ground, maintaining that analyses of justice must be realistic, that is, realizable under the best of foreseeable conditions. Against Rawls and others, this article argues that feasibility does not constrain what can count as an adequate principle of justice but nevertheless maintains that there are limits on such principles that derive in part from human nature, which divergent theories of justice must respect. It also distinguishes between different levels of analysis, some of which are governed by feasibility constraints.

What constraints do historical circumstance, institutional design and human nature place on the analysis of ideals such as justice? To what extent should these analyses be governed by judgements about the feasibility or otherwise of implementation, and judgements about what human beings are like? These questions attract different answers and give rise to competing conceptions of the nature and role of political theory. At one end of the spectrum there are those, like John Dunn, who argue that the analysis of political ideals such as justice should always be informed by a clear understanding of what is feasible given the social and historical forces that confront us and by an appreciation of the potential difficulties of institutional design.[1] At the opposite end of this spectrum, there are others, such as G. A. Cohen, who maintain that fundamental principles of justice are logically independent of issues of feasibility and questions about human nature.[2] According to the former position, if political theorists are to understand what justice means, they must possess considerable knowledge of the workings of our existing social, political and economic institutions as well as the context in which they function, whereas according to the latter position, these matters have no bearing on the nature of justice itself.

In between the two extreme positions I have sketched lie many political theorists, such as John Rawls[3] and Thomas Nagel,[4] who argue that the analysis of political ideals such

* Department of Politics, University of Southampton. The author would like to thank Chris Brown, G. A. Cohen, Cécile Fabre, John Horton, Ken Newton, David Owen, Graham Smith and Albert Weale for helpful comments.

[1] See J. Dunn, *Interpreting Political Responsibility* (Cambridge: Cambridge University Press, 1990), chap. 12, pp. 193–215.

[2] See G. A. Cohen, 'Facts and Principles', *Philosophy and Public Affairs*, 31 (2003), 211–45, and his draft manuscript 'Rescuing Justice from Constructivism'. I benefited from reading these pieces before writing the final draft of the present article. They cover some of the same ground and contain more extensive arguments for some of the conclusions that I had reached independently.

[3] See, especially, J. Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971) or J. Rawls, *A Theory of Justice*, revised edn (Oxford: Oxford University Press, 1999) (in the citations that follow, references will be given in the form p. x/y, where 'x' denotes the page number in the 1971 edition and 'y' denotes the page number in the 1999 edition); J. Rawls, *Political Liberalism*, paperback edn (New York: Columbia University Press, 1996).

[4] See T. Nagel, *Equality and Partiality* (Oxford: Oxford University Press, 1991).

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

as justice is governed by some considerations of feasibility and involves making some assumptions about human nature, but that it is nevertheless free to ignore a range of other questions about what can be realized here and now and what human beings are like. Consider, for example, the account of justice that Rawls defends in *Political Liberalism* and the articles leading up to this book. Rawls places a number of constraints on an adequate theory of justice, including the requirement that it must be realistic.[5] Indeed Rawls's official story of why he was led to abandon parts of the account he developed in *A Theory of Justice* maintains that the idea of a well-ordered society it presents is unrealistic 'because it is inconsistent with realising its own principles under the best of foreseeable conditions'.[6] In Rawls's view, a conception of justice is realistic only if it is capable of forming the basis of a stable society, and indeed one that is stable for the right reasons.[7] If a democratic society is to be stable, it must be such that it can be 'willingly and freely supported by at least a substantial majority of its politically active citizens'.[8] According to Rawls, a realistic account of justice is not obliged to take into consideration each and every historically contingent reason why it might fail to receive widespread support. But an account of justice is realistic for democratic regimes only if 'under the best of foreseeable conditions' it is capable of being the object of an overlapping consensus amongst reasonable persons without the oppressive use of state power.[9]

I propose to explore the issue of what constraints we should impose on the analysis of political ideals, focusing on the concept of justice. Despite the temptation to steer a middle path between the view that analyses of justice can legitimately ignore questions of feasibility and the view that these analyses are adequate only if they involve a careful investigation of the various constraints that human nature, historical context and institutional design place upon implementation, I shall try to show that there is a need for a kind of abstract theorizing about justice which is unconstrained by considerations of feasibility, although it has to be sensitive to at least some considerations that derive from human nature. I shall argue that abstract theorizing of this sort is a necessary complement to empirically grounded research that confronts more concrete questions about what is possible for us here and now.

The article is structured as follows. In the first four sections I shall present the arguments for imposing a variety of different constraints on the analysis of the concept of justice: in the first three sections I explore some limits that might be thought to arise from considerations of feasibility, stability or human nature; in the fourth section I examine the idea that a principle of justice is adequate only if its application is publicly checkable. In the fifth section, I shall outline a conception of political theory that involves different levels of analysis, each level introducing further constraints.

[5] See Rawls, *Political Liberalism*, p. xix, pp. 38–9. This is not the only possible reading, however. Rather than placing a constraint on any adequate theory of justice, it might seem that Rawls is merely arguing that it is desirable for a theory of justice to be realistic. According to this interpretation, Rawls is maintaining that it is better for a theory of justice to be realistic, but allowing that it is possible in principle for an adequate theory of justice, or indeed the best theory of justice, to be unrealistic. For a careful discussion of Rawls's position on these matters that is broadly consistent with my interpretation, see D. Copp, 'Pluralism and Stability in Liberal Theory', *Journal of Political Philosophy*, 4 (1996), 191–206.

[6] Rawls, *Political Liberalism*, p. xix.

[7] Rawls, *Political Liberalism*, p. xlii.

[8] Rawls, *Political Liberalism*, p. 38.

[9] See J. Rawls, *Justice as Fairness: A Restatement* (Cambridge, Mass.: Harvard University Press, 2001), p. 4.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

HISTORICAL LIMITS

The abstract analyses of ideals favoured by many contemporary political philosophers are sometimes criticized on the grounds that they are *politically irrelevant*. To the extent that these analyses ignore the issue of what opportunities exist for implementing the visions they offer, they run the risk of being irrelevant to our political activity. Presenting a different approach, John Dunn argues that '[t]he purpose of political theory is to diagnose practical predicaments and to show us how best to confront them'.[10] This, he argues, requires political theory to be conducted in the light of a proper understanding of the social, political and economic setting of our lives and the possibilities for change that it affords: 'political theory has no choice but to try to tell us how to act, given that this indeed is where we now are'.[11] Those who conceive political philosophy as 'the application of ethics to political topics', and who embark upon the conceptual analysis of ideals without considering how these ideals might be realized in the social, political and economic context we confront, are involved in an activity that is inherently limited. Their analyses will possess only marginal political relevance, for they will have no definite practical implications in the absence of an understanding of the causal forces that bear upon implementation.[12]

In presenting this vision of political theory Dunn is not denying the need to reflect upon our ideals or to be morally imaginative. He maintains that political theory requires some account of the direction in which we could coherently and justifiably wish the world to change, but argues that this is secondary to 'the causally adequate analysis of the human world in terms of its openness to human betterment'.[13] The analysis of political ideals needs to be grounded in a causally adequate understanding partly in order to counteract the natural tendency for *wishful thinking*, which can manifest itself as a stubborn belief that the world must be such that it can be shaped by one's ideals and is really nothing more than the belief that change must be possible because it is desirable. Forcing political theorists 'to locate the levels of moral ambition which they espouse within their best causal understanding of the human world as this is … precludes them in consequence from subordinating their understanding of how it really is to the importunities of their own projective desires'.[14]

Are these arguments against 'ungrounded' political theory effective? The claim that political theory will be irrelevant to our political activity unless it starts from a rigorous account of what is possible for us, situated as we are, invites the response: 'Why must political theory always be directly relevant for political activity?' Why should we suppose that the *sole* purpose of political theory is 'to diagnose practical predicaments and to show us how best to confront them'? Of course, those who deny that political theory must be directly relevant to politics are not thereby committed to denying that political theory can contribute to the important and valuable task of identifying and evaluating feasible alternatives. The point is merely that its remit is broader than this.

Indeed, if political theory is constrained to start from the identification of what is possible, it becomes vulnerable to 'adaptive preference formation'. G. A. Cohen makes this point in the context of discussing the ways in which socialists have come to terms with the thought that in the present circumstances the best that is possible is some

[10] Dunn, *Interpreting Political Responsibility*, p. 193.
[11] Dunn, *Interpreting Political Responsibility*, p. 196.
[12] Dunn, *Interpreting Political Responsibility*, pp. 194–5.
[13] Dunn, *Interpreting Political Responsibility*, p. 196.
[14] Dunn, *Interpreting Political Responsibility*, p. 196.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

form of capitalism, rather than true socialism which requires equality of a kind that is unachievable within a capitalist system.[15] Socialists who accept this analysis do not have reason to be fully satisfied with the best possible form of capitalism. If they were fully satisfied with it, their reasoning would involve adaptive preference formation. They would in effect be deciding that there is no reason to want 'true socialism' because it is impossible to achieve here and now: perhaps they change their evaluative criteria or perhaps they change their view of how well true socialism measures up to those criteria, but they do so merely because of its unavailability.[16] (Compare the fox in Aesop's fable, who decides that he does not want the grapes which are out of his reach because they cannot really be sweet and juicy.[17])

In effect, Dunn's charge that political theorists who work out their visions in the absence of an adequate causal understanding of the world are in danger of wishful thinking is met by the counter-charge that unless we begin from minimal constraints and ignore questions of feasibility we shall be prone to adaptive preference formation. In response, he can of course point out that those who start from what is possible here and now, and limit their judgements about what is ideal to this range, need not be guilty of adaptive preference formation. They are concerned with the question: 'what is ideal under present historical circumstances?' not, say, with the more abstract question: 'what is ideal under optimal conceivable conditions?' This response is fine as far as it goes. But in order to answer the former question properly, theorists will need to understand *how* the considerations they cite as reasons for thinking that some set of arrangements are ideal under present historical circumstances justify that conclusion. That will require some understanding of what weight (if any) those considerations would have in different circumstances when other considerations also enter the picture and need to be taken into account. This may not require us to address the question 'What is ideal under optimal conceivable conditions?', but it does require us to abstract from the particular set of circumstances which limit current possibility, otherwise there is the danger that our ideals will be inadequately rather than better grounded.

Starting from judgements about what is feasible may also mire us in difficulties that could be avoided were we to start from more minimal empirical constraints. Our understanding of the causal forces that bear upon our current circumstances is bound to be limited and often contested. Disagreement about the nature and scope of these forces means that there will be no shared basis from which to judge what is feasible. Setting to one side the issue of what is feasible is a way of bracketing these controversies. Even where there is agreement about the nature and scope of the causal forces that govern our current predicament, it may rest upon a mistaken analysis of our circumstances. As a result, limiting ourselves by what we judge to be possible in the foreseeable future may make us ill-equipped for the changes that actually occur, for these are sometimes more dramatic than we envisage. If our ideals are constrained by what we judge to be possible, now or in the foreseeable future, we will have nothing to fall back on if it turns out that these judgements are mistaken. Given our fallibility and the limits on what we can reliably predict, we need to explore ideals in a way that is not too tightly constrained by what we

[15] G. A. Cohen, *Self-Ownership, Freedom, and Equality* (Cambridge: Cambridge University Press, 1995), p. 254.

[16] See Cohen, *Self-Ownership, Freedom, and Equality*, pp. 254–7.

[17] See J. Elster, *Sour Grapes* (Cambridge: Cambridge University Press, 1983), chap. III.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

judge to be possible, in order to prevent our political thinking from being rendered irrelevant by changes we did not, or could not, foresee.

The arguments I have been giving do not detract from the importance of carefully investigating and determining (in so far as that is possible) the feasibility of implementing an ideal. Unless attention is paid to this question, attempts at implementation may be counter-productive or have other undesirable consequences. Realists in International Relations have insisted that attempts by statesmen to pursue ideals of justice in the international arena will make the world more unjust or more unstable, or both.[18] Though these fears are often over-stated, it is hard to deny that the history of intervention provides some rational basis for them. Furthermore, as Robert Goodin has emphasized, in situations where an ideal cannot be fully realized, the second best may turn out to be rather different.[19] If we try to get as close to the ideal as possible, then we may end up creating a state of affairs that is worse than what would be achieved by attempting to implement the second best. These observations show the importance of investigating feasibility constraints. Unless these constraints are correctly identified, important consequences and costs of implementing a proposed ideal will be ignored and these are clearly relevant to determining whether it *should* be implemented.[20] But my question is not: why should we think that it is important to investigate these constraints? It is rather: what reason do we have for thinking that any adequate analysis of an ideal such as justice must be conducted in the light of an investigation of what is feasible? Is there some conceptual incoherence involved in saying: 'This is what justice involves, but there is no way it could be implemented'? As yet, no reason has been found to answer this question in the affirmative. In the next section I shall explore it in more depth by considering how problems that may arise in designing institutions affect the issue of whether an account of justice is adequate or not.

LIMITS IMPOSED BY PROBLEMS OF INSTITUTIONAL DESIGN

In their *Companion to Contemporary Political Philosophy*, Robert Goodin and Philip Pettit maintain that the concern of political philosophy is 'to identify the sort of political institutions that we should have'.[21] This places the design of political institutions at the heart of political philosophy in a way that will be contested by those who think that its primary concern should be to analyse political ideals and that this can be done without considering how these ideals are to be institutionalized. But Goodin and Pettit's account of political philosophy does not, by itself, entail that we cannot adequately analyse ideals without considering problems of institutional design. It is undeniable that problems of implementation need to be considered in answering the question of whether justice (conceived in some way that is being proposed) should be implemented, but the real issue

[18] See, for example, H. Morgenthau, *In Defence of the National Interest: A Critical Examination of American Foreign Policy* (Washington, D.C.: University Press of America, 1982), especially pp. 38–9; G. Kennan, *American Diplomacy 1900–1950* (Chicago: University of Chicago Press, 1950), p. 101.

[19] See R. Goodin, 'Designing Constitutions: The Political Constitution of a Mixed Commonwealth', *Political Studies,* 44 (1996), 635–46; R. Goodin, 'Political Ideals and Political Practice', *British Journal of Political Science*, 25 (1995), 37–56. These issues are discussed further in the fifth section below.

[20] See J. Carens, 'Compensatory Justice and Social Institutions', *Economics and Philosophy*, 1 (1985), 39–67, p. 39.

[21] R. E. Goodin and P. Pettit, eds, *A Companion to Contemporary Political Philosophy* (Oxford: Blackwell, 1993), p. 1.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

is whether they can be bracketed in answering the more abstract question of whether this is the best way of conceiving 'justice'.

It might be argued that if the costs of implementing a proposed ideal are sufficiently high, or if the proposed ideal is impossible to institutionalize, then it cannot legitimately be regarded as an ideal at all, so problems of implementation cannot be bracketed in analysing ideals. Consequentialist moral theorists, for example, might argue that any principle has to be assessed in terms of whether its implementation would maximize overall welfare or well-being. According to this view, whether a principle of justice is acceptable or not depends on whether its realization would tend to maximize welfare or well-being summed across society, with each person's welfare or well-being counting equally. If this is so, the acceptability of a principle of justice depends upon the consequences of implementing it; therefore, no principle can legitimately be regarded as acceptable in the absence of some consideration of how it might be institutionalized.

But this argument applies only to consequentialist theories, and only to some of them. Those who reject the idea that the consequences of institutions can all be compared (in terms of their effects on welfare, well-being or whatever), perhaps because they believe that there is a plurality of incommensurable values, would not be sympathetic to it. Even some consequentialists have reason to reject the argument, for example, those who accept that there is a plurality of values but allow that these can be compared sufficiently to make meaningful on balance judgements about what policies, practices or institutions can be expected to produce the most overall value. Defenders of a sophisticated form of consequentialism of this kind can allow that the cost of implementing an ideal may be high, in terms of its impact upon other values, without that undermining its status as an ideal.[22]

There is, however, a different argument to explain why issues of institutional design cannot coherently be ignored, which has relevance for consequentialists and non-consequentialists alike. This argument appeals to the principle that 'ought implies can' which is common to both traditions. If 'ought implies can', it would seem to follow that if it is impossible to devise an institution to track the requirements of some proposed analysis of justice, it makes no sense to say that these requirements ought to be tracked. But if it makes no sense to say that they ought to be tracked, then it cannot represent a failure of justice not to do so. Arguing along these lines, Andrew Levine rejects the idea that it is unjust not to reward effort, on the grounds that 'existing structures reward effort poorly, if at all, and … because it could hardly be otherwise, there are no compelling reasons of justice that should lead us to regret this state of affairs'.[23] Joseph Carens maintains that the principle that ought implies can does not undermine an ideal which cannot be implemented because of contingent political, social or economic circumstances, but argues that it does subvert ideals which are 'inherently unworkable': 'if someone puts forward a principle of distribution but cannot describe, even in theory, how to arrange social institutions to implement that principle in a stable and effective way, then we are bound to ask whether the principle can really be regarded as an ideal for human societies'.[24]

[22] Consequentialism is a diverse tradition. See J. Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1986), chap. 11, especially section 1, for ways of understanding its diversity.

[23] A. Levine, 'Rewarding Effort', *Journal of Political Philosophy*, 7 (1999), 404–18, p. 417. Consider also Alexander Rosenberg's claim that 'if equality is an unattainable goal, then it cannot be a morally obligatory one'. So by arguing that equality is an unattainable goal, Rosenberg takes himself to have shown that 'a commitment to equality is not morally obligatory'. See A. Rosenberg, 'Equality, Sufficiency, and Opportunity in the Just Society', in E. F. Paul, F. D. Miller and J. Paul, eds, *The Just Society* (Cambridge: Cambridge University Press, 1995), pp. 54–71, at p. 55.

[24] Carens, 'Compensatory Justice and Social Institutions', p. 64.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

The idea that an account of justice is inadequate if it cannot be institutionalized because 'ought implies can' rests on the supposition that principles of justice can be expressed in terms of 'ought' statements. If principles of justice can be expressed in terms of ought statements, and ought implies can, then it would seem to follow that any adequate principle of justice must be capable of implementation.[25] But this argument ignores a distinction that needs to be drawn between a concrete use of the term 'ought', which is governed by what is possible and which licenses the inference from 'ought' to 'can', and an abstract use of the term, which brackets issues to do with what is possible and which does not license this inference. It is the abstract use of 'ought' that is involved in principles of justice, and indeed in other general moral principles, such as the principle that people should keep their promises. This abstract sense may also be used in particular judgements when these are intended to draw attention to some morally relevant feature of a situation, rather than express an all things considered judgement about what should be done. When this is the case, 'A ought to do X' means something like, 'A has a moral reason to do X'.[26] So, for example, employing the abstract sense of ought, we might say 'Pauline ought to have visited her friend because she promised', even though we know she was unable to do so because, say, there was a train derailment. Here we mean that Pauline had a moral reason to visit her friend, a reason which she would have needed to take into account had she been able to visit the friend.

This distinction between the abstract sense of 'ought' and its concrete sense is also relevant in another way. Not only does the abstract sense of 'ought' bracket issues to do with what is feasible, it brackets other considerations that are relevant to the determination of action or policy as well. For example, it permits the claim that even though Pauline ought to keep her promises (for there is always a reason for her to do so), there are occasions on which, all things considered, she ought not to keep a promise because that would require her to violate some more important obligation. In a similar way it permits the claim that there are circumstances when we ought not to establish just institutions because of the costs in terms of other values of doing so. Some might reject this latter claim on the grounds that it is inconsistent with regarding justice as the first virtue of social institutions. Rawls, for example, regards justice in this way and takes it to mean that if an institution or law is unjust, then it ought to be reformed.[27] But I think that Rawls's interpretation of the idea that justice is the first virtue of social institutions is problematic. He concedes that his initial formulation of this idea will require refinement, ignoring as it does all of the potential costs of reform, some assessment of which must surely be made in addressing the issue of whether an unjust institution ought to be reformed.[28] He allows that injustice can legitimately be tolerated but only if the alternative would mean greater injustice.[29] That still seems too restrictive, however, making justice too uncompromising: if reforming an unjust institution would require massive sacrifice of other values, then there may be good reason to stick with the status quo. So I think we should understand the claim that justice is the first virtue of institutions in such a way that it attributes considerable weight to justice but nevertheless allows that it can be traded against other values. The principle that if an

[25] See Cohen, 'Facts and Principles', section M.

[26] Some would argue that the abstract use of the term 'ought' is never legitimate, on the grounds that whether a property or a feature of a situation is morally relevant (and, if so, in what way) always depends upon what other properties or features are present. See J. Dancy, *Moral Reasons* (Oxford: Blackwell, 1993).

[27] See Rawls, *A Theory of Justice*, pp. 3–4/3–4.

[28] See Rawls, *A Theory of Justice*, p. 4/4.

[29] See Rawls, *A Theory of Justice*, p. 4/4.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

institution is unjust, it ought to be reformed is much too simplistic. A more refined formulation of it would accept that a variety of different sorts of costs of reforming an unjust institution need to be taken into account in determining whether it ought to be reformed.[30]

I have claimed that the principle that 'ought implies can' does not show that the practice of analysing ideals without considering issues of implementation is flawed, because there is an abstract use of 'ought' which brackets the question of what is feasible. But Carens thinks that there are more serious difficulties with the practice, difficulties that are independent of the principle that 'ought implies can'. He maintains that in some cases we cannot even *fully understand* an ideal without confronting problems of implementation.[31] He provides a powerful illustration of his argument by considering the ideal of compensatory justice: 'the view that inequalities of income should be permitted only to the extent that they compensate for differences in the advantages and disadvantages of different sorts of work'.[32]

In the course of clarifying this ideal, Carens draws attention to a point which he believes any analysis of it should take into account, namely, that there are differences in the burdens that people bear in their work, yet what feels burdensome to one person may not to another.[33] He argues that an approach which appeals to the relative disutilities of different forms of work might seem to be on the right track but runs into a number of difficulties. For example, such an approach to compensation would appear to entail paying different rates to two people who perform the same job just because one likes it whereas the other hates it. Carens then argues that the only way of respecting our intuitions about the burdens of work, whilst avoiding the problems a relative utility approach faces in providing just compensation for them, is to look at the issue of institutionalization. When we do that we will see that the best model of compensatory justice is provided by a market which operates under the assumption of perfect competition and in which suppliers of labour services have identical skills. Under such a scheme, 'jobs that most people found onerous and unpleasant would be relatively well paid and jobs that most people found pleasant and interesting would be relatively poorly paid. Yet no one would be paid more than another for doing the same job simply because he hated it whereas the other liked it'. So, Carens concludes, 'the best way to explain the ideal of compensatory justice is simply to point to the intuitive appeal of the way the market works under [conditions of perfect competition and identical skills]'.[34] He then goes on to argue that the problems faced in implementing the ideal stem from the fact that in the real world markets do not operate under these conditions: there are considerable differences in natural talents and the skills people are able to acquire and competition is not perfect in the economist's sense.

But it is far from clear that Carens's argument is successful in showing what he thinks it does, namely, that sometimes we cannot fully understand an ideal unless we consider how it might be implemented. For arguably the ideal of compensatory justice is fully clarified simply by making the points that he does concerning its implications for different sorts of cases, for example, that when one person does a job that is more unpleasant or onerous than the job another does, the former should receive a larger income; that when two people do the same job the mere fact that one likes it whereas the other does

[30] See Cohen, 'Rescuing Justice from Constructivism', for further discussion.
[31] Carens, 'Compensatory Justice and Social Institutions', p. 39.
[32] Carens, 'Compensatory Justice and Social Institutions', p. 39.
[33] Carens, 'Compensatory Justice and Social Institutions', p. 46.
[34] Carens, 'Compensatory Justice and Social Institutions', p. 55

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

not should make no difference to their incomes.[35] Indeed it is a prior understanding of the ideal revealed in judgements of this kind which allows us to assess how well particular schemes would institutionalize the ideal. Carens operates with that prior understanding throughout his article, and indeed it plays a role in his argument in the final section that 'movement in the direction of equality is the best that we can hope for if we are committed to the ideal of compensatory justice'.[36] Raising problems of institutional design may help to tease out some of the logical implications of the ideal, but in principle at least these can be drawn out without doing so.

LIMITS IMPOSED BY STABILITY AND HUMAN NATURE

Human nature is generally regarded as another potential source of constraints on any adequate conception of justice. It is true that conceptions of justice must take a stand on whether, for example, justice requires the promotion of basic human interests. Those who believe that it does will have to defend some account of these basic interests if they are to be in a position to apply this principle to the design of institutions. But the controversial issue of whether justice requires us to promote basic interests, which divides some theorists of justice, is different from the issue that concerns me here: namely, whether the effects of natural human dispositions on the stability or feasibility of social arrangements has any bearing on what can count as an adequate principle of justice.

Let us begin by considering the issue of stability. If a distributive principle would place large burdens on some people, making it hard for them to reconcile themselves to the institutions that this principle underwrites, does any resulting instability of these institutions by itself give us a reason (even if that reason is inconclusive) to reject it as a correct principle of justice? In *A Theory of Justice*, Rawls maintains that any adequate conception of justice must generate its own support, that people living in a society whose institutions are governed by it must tend over time to develop ways of thought and judgement which lead them to support that conception for its own sake and to accept the corresponding institutions.[37] And in *Political Liberalism*, he argues that any adequate conception of justice for a democratic regime must be realistic because only then is it capable of providing the right basis for stability.[38] He does not suppose that a conception of justice is unrealistic for a democratic society whenever there are obstacles in the way of its being widely accepted; an adequate conception of justice for such a society is not required to accommodate each and every historically contingent reason why it might fail to receive widespread support. For Rawls, what matters is whether a conception of justice can be realized 'under reasonably favourable but still possible historical conditions, conditions allowed by the laws and tendencies of the social world'.[39]

In Rawls's view, a conception of justice is realistic for a democratic regime (and will provide an appropriate basis for stability in it) only if, under the best of foreseeable conditions, it is capable of becoming the object of an overlapping consensus amongst reasonable persons without the oppressive use of state power, on the assumption that reasonable persons possess the desire to engage in fair co-operation with others and

[35] If judgements of this sort could not be made, then of course there would be a strong case for saying that we do not understand the ideal.

[36] Carens, 'Compensatory Justice and Social Institutions', p. 67

[37] Rawls, *A Theory of Justice*, pp. 138/119, 177/154; Rawls, *Justice as Fairness: A Restatement*, pp. 124–5.

[38] See Rawls, *Political Liberalism*, p. xix; p. 38; p. xliii.

[39] Rawls, *Justice as Fairness*, p. 4.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

acknowledge the existence of burdens of judgement which make possible reasonable disagreement over the acceptability of comprehensive moral doctrines.[40] This in effect places a stability constraint on an adequate conception of justice, one which rules out the possibility of a conception of justice being adequate for a democratic regime when its defence appeals to the truth of any comprehensive moral doctrine that is subject to reasonable disagreement, and which is therefore unlikely to provide the basis of a free and reasoned overlapping consensus.

In response to Rawls's account of the relationship between justice and stability, we should pose the following question. Is it really *conceptually incoherent* to suppose that a just society might be inherently unstable because human beings, as a result of their natures, would find it hard to live under such arrangements? For there seems no reason to think that unstable institutions are necessarily unjust, nor for supposing that there is any conceptual incoherence in any of the following ideas: that one set of institutions might be more stable but less just than an alternative set; that a stable but to some degree unjust state might be more desirable than a fully just but deeply unstable state; that 'justice may…be fragile even in the best of circumstances'.[41] In short, we have been given no reason to deny that justice and stability are independent values (which, if they conflict, require a proper balance to be struck between them). Of course, when just institutions are stable, they are more likely to endure over time, which is itself better from the point of view of justice, and this provides part of the reason for valuing stability. But this is not to claim that stable institutions are in themselves more just, nor to deny that stability may be valued for other reasons. To maintain that it has independent value is not necessarily to claim that it has non-instrumental value, however. We could in principle treat justice and stability as independent values whilst arguing that, say, stability is independently valuable only because it promotes personal autonomy by making it easier for individuals to form and pursue their own conceptions of the good. But if the two values are genuinely independent, the stability or otherwise of the institutions underwritten by a set of principles of justice has no bearing on the correctness of these principles *as principles of justice*.[42] Here I am not denying that there may be other reasons for following Rawls in avoiding an appeal to comprehensive moral doctrines in developing a conception of justice. For example, it might be argued that it is important for principles of justice to be publicly justifiable, because this is required by respect for citizens, or for a regime based upon these principles to be legitimate. But these reasons are independent of a concern with stability.

Some might claim that even if justice and stability are independent values, human nature always constrains what can count as an adequate principle of justice because no principle of justice could be adequate if it would make unreasonable demands upon some. This proposal would not generate a *feasibility* constraint on principles of justice, but its application would require making some assumptions about human nature, for what counts as an unreasonable demand will be determined in part by what human beings are like. But is the question of whether such a constraint can be justified any different in kind from a range of other questions about what justice requires that may also be informed by assumptions about human nature and which were bracketed at the beginning of this section? For example, is it any different in kind from the question of whether justice requires the state to meet the basic needs of its citizens or promote their basic interests,

[40] See Copp, 'Pluralism and Stability in Liberal Theory', pp. 196–201.
[41] Copp, 'Pluralism and Stability in Liberal Theory', p. 204.
[42] See Cohen, 'Rescuing Justice from Constructivism', for further discussion.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

which are matters of dispute between different theories of justice? My hope is that by exploring the idea that no principle of justice could be adequate if it would make unreasonable demands upon some, we will be able to justify a constraint on principles of justice which is more abstract than, say, the idea that no principle of justice could be adequate if it would leave some unable to meet their basic needs, and which is relatively uncontroversial since it is independent of disputes between particular theories of justice such as those defended by libertarians and left-liberals.

From the perspective of some theories of justice, in particular those influenced by T. M. Scanlon's work in moral theory, no adequate principle of justice could make unreasonable demands on the individuals subject to it, since the test of a correct principle of justice is precisely whether any one could reasonably reject it, and a principle that made unreasonable demands on some would fail that test.[43] But is there a constraint of this kind that even those who are not persuaded by a Scanlonian approach to theories of justice ought to accept? If there is to be such a constraint, some account will need to be given of what makes a demand unreasonable that does not simply appeal to prior principles of justice. For if we say that a demand is unreasonable only if it is inconsistent with the demands which the correct principles of justice would make, the condition I am envisaging could not place a constraint on what can count as an adequate principle of justice.

In addressing this question it is instructive to consider an objection that is sometimes made against act consequentialism, namely, that it is unreasonably demanding because it requires individuals to maximize the good.[44] When stated bluntly, this objection to act consequentialism invites the response that morality just is demanding; the demandingness of a principle is not, by itself, sufficient to justify its rejection. But the objection can be developed so that it is not vulnerable to this rejoinder. It is not simply that act consequentialism is demanding; it is demanding in a way that would alienate people from their commitments.[45] Whenever it was possible for a person to increase the overall amount of good in the world, even by some small amount, they would be morally obliged to put aside their own commitments in order to do so.[46] Its demandingness threatens the very possibility of a person's *having* commitments (other than a commitment to maximize the good)[47] and therefore of being an agent at all.

How should we understand this objection to act utilitarianism? It might be argued that it reveals a clash of values between the maximization of the good and integrity, and that individuals need to balance their pursuit of the good against the threat that this may pose to their own integrity. But this would be to misunderstand the lesson of this argument. The argument is not pitting some independent value, integrity, against the maximization of the good; instead it is implicitly appealing to a constraint that any adequate moral theory, including act consequentialism, must meet, namely that such a theory must

[43] See B. Barry, *Justice as Impartiality* (Oxford: Oxford University Press, 1995). Scanlon himself is reticent about the extent to which his contractualist theory applies to principles of justice. He maintains that his account is intended to apply only to part of morality, namely that concerned with our duties to others: see T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass.: Harvard University Press, 1998), pp. 6–7.

[44] For a recent discussion of this issue, see B. Hooker, *Ideal Code, Real World: A Rule Consequentialist Theory of Morality* (Oxford: Oxford University Press, 2000), pp. 149–58.

[45] For the classic source of this objection, see B. Williams, 'A Critique of Utilitarianism' in J. J. C. Smart and B. Williams, *Utilitarianism: For and Against* (Cambridge: Cambridge University Press, 1973), pp. 77–150, at pp. 108–18.

[46] See Hooker, *Ideal Code, Real World*, pp. 150–1.

[47] See Hooker, *Ideal Code, Real World*, pp. 155.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

not make unreasonable or excessive demands on individuals of a kind which would undermine the very possibility of their having commitments. We might suppose that principles of justice are subject to a similar constraint. To show that a principle of justice made unreasonable demands of the relevant sort, we would need to show that obliging agents to live under institutions that embody the principle, and to follow that principle in the appropriate way in their lives, would, for some, threaten the very possibility of having commitments. Arguments of this kind are sometimes available. Consider, for example, Ronald Dworkin's objection to a particular interpretation of equality of resources, namely that it would lead to the slavery of the talented, for the talented would come to own very few hours of their own labour time.[48] The details of this objection, and the conception of equality of resources to which it is supposed to apply, do not matter for my present purposes, nor whether its conclusion is justified; the point is simply that it has the right character, for if it were successful, it would show that such a conception of equality would be unreasonably demanding in at least one way because for some (that is, the naturally talented) it would threaten the very possibility of having commitments.

Again, this kind of argument is not pitting the value of integrity against that of justice; instead it is appealing to the idea that any adequate principle of justice must allow space for each individual to have commitments that are not exhausted by their commitment to that principle or to the institutions which embody it. An argument of this character has its limits, however. To show that a principle of justice falls foul of the relevant constraint, it is not enough to show that, for instance, it would require some to lead a life of poverty, or to show that it would forbid some from practising their own religion, or that more generally it would require some to make considerable personal sacrifices. These consequences might indeed undermine the possibility of a person's leading his own life in his own way, of leading a life that is autonomous (and some might reject the principle because they ground their theory of justice in the importance of each person's autonomy) but they would not undermine the possibility of his having commitments. Principles of justice that would legitimate societies in which some live in affluence and others in poverty, or which sanction illiberal restrictions on religious practice, or which require the talented to forgo the receipt of benefits for the exercise of those talents, cannot be defeated by the constraint I am defending merely by showing that they have results which some will find hard to accept. They would fail to meet this constraint only if it can be shown that the burdens they impose are unreasonable in at least one way because they make it impossible for some to have commitments.

I do not mean to suggest that the only way in which a burden could be unreasonable for a person is by making it impossible for him or her to have any commitments. There may be other ways in which a burden can be unreasonable, and indeed other accounts of what make it unreasonable that do not simply appeal to a principle of justice in determining what is to count as unreasonable. I have merely defended one such account and the constraint that follows from it, which, it seems to me, rival theorists of justice, whatever their approach, should find acceptable.

OTHER CONSTRAINTS ON ANALYSES OF THE CONCEPT OF JUSTICE

Theorists have proposed a number of constraints on principles of justice that they have

[48] See R. Dworkin, *Sovereign Virtue: The Theory and Practice of Equality* (Cambridge, Mass.: Harvard University Press, 2000), p. 90

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

claimed are independent of considerations of feasibility and independent of whether these principles make unreasonable demands on the individuals required to live under them. Some of these constraints derive from substantive judgements about what justice requires that are entrenched in our practices and which we cannot easily imagine abandoning, for example, the judgement that racial discrimination is unjust, or that slavery is unjust. These are what Rawls calls considered judgements, that is, 'provisional fixed points which we presume any conception of justice must fit',[49] 'judgments in which our moral capacities are most likely to be displayed without distortion'.[50] Rawls allows that our considered judgements may sometimes be abandoned in the light of reflection – that is implicit in describing them as *provisional* fixed points – but argues that they provide a way of testing a theory of justice, as part of a process designed to achieve a 'reflective equilibrium' between the theory and these considered judgements.

Other potential constraints on the concept of justice emerge not from prior judgements about what is just or unjust but rather from general constraints on what can count as an adequate principle of justice of a sort that I have not so far considered. For example, some have proposed that the application of any adequate principle of justice must be publicly checkable. Rawls maintains that principles of justice apply to the basic structure of society, most plausibly understood as comprising all social institutions that have profound effects on the individuals subject to them. Andrew Williams argues that this has important implications for what can count as an adequate principle of justice. For Rawls defines an institution as a public system of rules,[51] and a necessary condition of a system of rules being public is that it should be possible for others to check the extent to which a person's behaviour conforms to these rules.

Williams maintains that the requirement of public checkability means that rules have to be capable of formulation in ways that are not too vague or ambiguous, and that the information required in order to determine whether these rules have been obeyed should be available to others.[52] He argues that this condition rules out the possibility that a principle of justice might constrain an individual's economic behaviour in some contexts, for example his pursuit of his own interest in wage negotiations or career choices. In Williams's view, there is no plausible rule that clearly and unambiguously tells a person how much weight to give to her own self-interest in such decisions. Even if such a rule could be formulated, its application would not be publicly checkable, given the difficulties of obtaining reliable information about how burdensome different jobs are for different individuals. Williams expresses the requirement of publicity in terms of the slogan 'justice must be seen in order to be done'.[53]

But why should we regard the idea that the application of principles of justice must be publicly checkable as an internal constraint on these principles rather than as an appeal to an independent value, that of publicity? Is this a genuine constraint on what can count as an adequate conception of justice, or does it illustrate instead the way in which justice

[49] Rawls, *A Theory of Justice*, p. 20/18.
[50] Rawls, *A Theory of Justice*, p. 47/42.
[51] See Rawls, *A Theory of Justice*, p. 55/47.
[52] See A. Williams, 'Incentives, Inequality, and Publicity', *Philosophy and Public Affairs*, 27 (1998), 226–48, pp. 233–4.
[53] Williams, 'Incentives, Inequality, and Publicity', p. 247.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

can come into conflict with *other* values? If the latter, the value of publicity will need to be taken into account when designing institutions, in order to produce a proper balance between it and justice, but it does not constrain what can count as an adequate conception of justice.

I shall refrain from making a judgement about whether the Rawls-inspired account of publicity given by Williams adequately captures the notion.[54] It seems to me that, to the extent it does capture it, the appeal of publicity derives from two sources, neither of which would justify the conclusion that it places a genuine constraint on what can count as an adequate principle of justice. The first source of its appeal is the idea that justice itself is better promoted by publicly checkable rules; the second source is the idea that publicity tends to promote stability. Let us consider each of these ideas in turn.

When others can check whether rules have been observed, justice is better promoted, for two reasons. First, a person may be willing to act justly only if they have the assurance that others will do so as well, and sometimes they can have that assurance only if they know that the behaviour of others can be checked to see whether it complies with what justice requires. Secondly, unless others can check whether a rule has been observed, it may be possible for individuals to benefit unfairly from acts of deception without any chance of being found out; in this way publicly checkable rules reduce the incentives for deception. Both of these reasons for valuing publicly checkable rules appeal to the idea that such rules better promote justice itself. But it does not follow from them that the application of any adequate principle of justice must be publicly checkable. And it is hard to see what could justify that claim. In this context consider the distribution of domestic responsibilities, such as childcare and housework. It is not *incoherent* to argue that principles of justice apply to the distribution of these responsibilities, whilst at the same time conceding that there is no reliable, publicly checkable way of determining whether these principles have been observed.[55] This would count against trying to institutionalize them but does not undermine the idea that the distribution of these responsibilities can be assessed from the standpoint of justice.

If publicity were valuable solely because of the role that it can play in promoting justice, then justice and publicity would not be independent values. However, the remaining part of the story of why publicity is valuable appeals to the importance of stability, broadly construed, and here there is a potential justification for maintaining that justice and publicity are independent values (though not for holding that publicity is of non-instrumental value). When the application of principles of justice is publicly checkable, and justice can therefore be seen to be done, one potential source of conflict and mistrust is removed, making social institutions more effective, stable and enduring, and thereby helping individuals to make plans and pursue their own conceptions of the good.

But is public checkability also part of the broader idea of public justifiability which some regard as partially constitutive of justice?[56] Public justifiability requires that people should be able to justify to each other the principles which govern their basic institutions and personal behaviour, but this does not entail that they should be in a position to justify to each other the particular distribution of benefits and burdens that results from the

[54] In 'Rescuing Justice from Constructivism', Cohen argues that publicity, so understood, demands more than can be justified.

[55] For further discussion, see Cohen, 'Rescuing Justice from Constructivism'.

[56] See, for example, Barry, *Justice as Impartiality*.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

the application of these principles, for example, what particular individuals enjoy which benefits.

LEVELS OF ANALYSIS

The discussion so far suggests that we would do well to distinguish a number of different levels of analysis of an ideal, and acknowledge the legitimacy of each. The first level is the most abstract and involves the elucidation of the best principles to express the ideal limited only by the constraint that these principles should not place unreasonable demands on individuals living under the institutions they govern, this constraint to be understood in the manner proposed earlier. The second level involves identifying the best principles to express the ideal limited this time by feasibility constraints as well, that is, by what is possible given our current historical circumstances, and/or by what we know about human nature (for example, that people will find institutions that embody these principles intolerable even if they are not unreasonably demanding in the sense explained earlier), and/or by constraints on institutional design. The third level involves two stages: first, identifying how the ideal should be balanced against others so as to ensure that due weight is given to each, representing this in terms of indifference curves; secondly, identifying how the ideal should be balanced against others in the light of any feasibility constraints.

Moving from the first level to the third level involves increasing the number of constraints that are taken into account. Onora O'Neill distinguishes between abstraction and idealization in a way that is relevant to understanding the relation between the different levels: abstraction 'is a matter of bracketing, but not of *denying*, predicates that are true of the matter under discussion'.[57] Idealization, in contrast, involves making claims that are, strictly speaking, false, as a way of simplifying an argument; for example, it might be assumed that all human beings have capacities for rational choice though in fact some do not. The different levels of analysis I have distinguished involve varying degrees of abstraction, for instance, the first level abstracts from various constraints that are considered at the other two levels, but do not necessarily involve idealization or increasing degrees of idealization.

The distinction I have drawn between the first and second levels does not coincide with that between ideal and non-ideal theory, at least not as this is commonly understood. As Rawls employs the distinction in *The Law of Peoples*, for example, ideal theory is subject to moderately strong feasibility constraints which require it to be realistic in the best of foreseeable conditions, whereas non-ideal theory is concerned either with problems of non-compliance, where actors behave unjustly, or with unfavourable historical, social or economic conditions.[58] My contrast between the first two levels of theorizing allows that there is a legitimate form of reasoning concerning justice which is not subject to feasibility constraints, and indeed allows that reasoning at this highest level of abstraction may also be concerned with what justice requires in response to past or present injustices.

The different levels distinguished are not intended to establish a division of labour between, say, political scientists and political theorists. In order for someone to

[57] O. O'Neill, *Towards Justice and Virtue: A Constructive Account of Practical Reasoning* (Cambridge: Cambridge University Press, 1996), p. 40, italics in original.

[58] See J. Rawls, *The Law of Peoples* (Cambridge, Mass.: Harvard University Press, 1999), esp. pp. 4–5, 11–23, 89–91.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

be able to work in a sophisticated way at the second and third levels, he or she will need to understand the various moves that are made at the first level, but it does not matter much what we call him or her. My point is simply that it is important to distinguish these levels, and to be aware of which level or levels a thinker is working at, otherwise conceptual mistakes will be made. For example, we may end up thinking that a conception of justice is necessarily flawed because it is unrealistic, whereas there is a level of thinking about justice where the question of whether it is realistic is irrelevant.

Nor are these levels intended to represent some necessary temporal order of research: there is no reason why detailed work at the first level has to be completed before moving to one of the other levels. Work at the second and third levels will always be vulnerable to arguments developed at the first level since these may undermine the principles whose application is being considered at those levels, but this does not provide an adequate reason for waiting until matters are settled at the first level, not least because there is no reasonable prospect of convergence there, nor indeed of an individual thinker reaching conclusions that require no adjustment or refinement in response to critical reaction.[59]

It would be a mistake to dismiss as utopian the most abstract level of theorizing I have distinguished. Thomas Nagel insists that any adequate theory of justice must give due weight to two different standpoints, the impersonal standpoint, which produces 'a powerful demand for universal impartiality and equality', and the personal standpoint, which 'gives rise to individualistic motives and requirements which present obstacles to the pursuit and realization of such ideals'.[60] For Nagel, a theory of justice is 'utopian in the pejorative sense if it describes a form of collective life that humans, or most humans, could not lead through any feasible process of social and moral development',[61] whereas a non-utopian solution requires a proper balance between considerations that derive from both personal and impersonal standpoints.[62] According to Nagel, '[i]f real people find it psychologically very difficult or even impossible to live as the theory requires, or to adopt the relevant institutions, that should carry some weight against the ideal'.[63] Without wishing to deny the very real problem that Nagel has identified (nor the importance of distinguishing between personal and impersonal standpoints), I think that his way of framing it is misguided in so far as it dismisses as utopian any theorizing that simply brackets the personal standpoint, and the motives to which it gives rise, when these are concerned with what individuals will find tolerable or what they will be motivated to comply with. Distinguishing between the different levels of theorizing allows us to acknowledge the value of such theorizing without denying the importance of other forms of theorizing which give full weight to these aspects of the personal standpoint and take full account of the dangers of trying to 'instantiate an ideal when it is not "motivationally possible" '.[64] Indeed, theorizing which brackets these aspects of the personal standpoint is important if we are to retain a clear view of what justice requires that does not simply capitulate to human nature in a way that Nagel himself would regard as unacceptable.[65]

[59] Although the claim stands in need of clarification, there is much to be said for the idea that key political concepts such as justice and democracy are essentially contestable: see W. B. Gallie, 'Essentially Contested Concepts', *Proceedings of the Aristotelian Society*, 56 (1955–56), 167–98; A. Mason, *Explaining Political Disagreement* (Cambridge: Cambridge University Press, 1993), chap. 2.

[60] Nagel, *Equality and Partiality*, p. 4.

[61] Nagel, *Equality and Partiality*, p. 6.

[62] Nagel, *Equality and Partiality*, p. 24.

[63] Nagel, *Equality and Partiality*, p. 21.

[64] See Nagel, *Equality and Partiality*, pp. 6–7.

[65] See Nagel, *Equality and Partiality*, p. 26.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

There is a deep difference between bracketing the issue of what individuals will find tolerable (or be able to comply with) and ignoring the issue of whether a principle places unreasonable demands on some people. The latter issue may itself require principles of justice to determine what is to count as an unreasonable demand, or (as earlier) appeal to a conception of what makes a demand unreasonable that is independent of such principles. Either way, it raises questions that cannot legitimately be put aside, even at the most abstract level of theorizing about justice. In Nagel's terms, we need to distinguish between different aspects of the personal standpoint. Some of these aspects can be bracketed in theorizing at the most abstract level, whereas others cannot. Any adequate theory of justice must be concerned with the reasonableness of the demands that it would make on people (though different theories may give different verdicts on whether a demand is reasonable or not) but may for some purposes legitimately ignore their ability to tolerate, and comply with, institutions informed by its principles, in so far as these issues do not bear on the reasonableness or otherwise of the demands the institutions make.

The different levels I have distinguished do not bear any simple relationship to each other; there is no formula for moving from one level to another. The theory of the second best precludes any straightforward inferences from either the first level to the second level, or from the second level to the third level. According to this theory, when ideal arrangements are unobtainable, the second best may be very different.[66] More generally, what is best under one set of conditions may not be a very good guide to what is best under some very different set of conditions. When a set of principles defended at the first level cannot be fully implemented, either because of historical circumstance, or because human beings will find it intolerable, or because of problems of institutional design, then the best that can be achieved in practice may turn out to be very different from what would be achieved simply by trying to implement this set of principles as far as possible, especially if we then move to the third level of analysis and take into account other ideals. Consider, for example, a conception of justice which would require us to track the choices and efforts that individuals make, rather than the circumstances against which they lead their lives. This analysis of justice at the first level may give way to a very different picture at the second and third levels. Given the difficulties of designing institutions that track choice and effort rather than circumstance, and given the costs in terms of other values in attempting to do so, it may be that the best we can achieve in practice is represented by an unconditional citizens' basic income.[67] And, of course, even this may be politically unrealistic.[68]

At what level does Rawls's theory of justice operate? It seems to function at all three levels.[69] His view that distributive shares should not be influenced by factors arbitrary from

[66] For a more precise statement of the theory, see R. G. Lipsey and K. J. Lancaster, 'The General Theory of the Second Best', *Review of Economic Studies*, 24 (1956), 11–33. For applications to politics, see Goodin, 'Political Ideals and Political Practice'; Goodin, 'Designing Constitutions: The Political Constitution of a Mixed Commonwealth'.

[67] See Carens, 'Compensatory Justice and Social Institutions', pp. 65–7, for an argument that the best way of implementing the ideal of compensatory justice would be to equalize incomes within the constraints of efficiency.

[68] Consider another example, furnished by Joseph Carens. He asks us to suppose that a just world order would allow free movement of people but argues that even if this is so, it does not follow that we should push for borders to be made as open as possible here and now. The second best may involve 'considerable closure so that other features of what a just world requires may be developed within a protected space' (J. Carens, 'Realistic and Idealistic Approaches to the Ethics of Migration', *International Migration Review*, 30 (1996), 156–70, p. 168).

[69] There is, however, the difficult issue (which Cohen in effect addresses in 'Rescuing Justice from Constructivism') of whether Rawls's constructivist method precludes Rawls from operating at the first level. For

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press

the moral point of view, and his (largely implicit) conception of what counts as such a factor, appears to be an example of analysis at the first level.[70] His argument that desert does not provide us with an adequate principle of justice, a 'non-arbitrary' factor to which distributions should be sensitive, because we cannot in practice distinguish between that part of a person's achievements which are due to factors beyond their control (such as their native endowments or their social circumstances) and that part which is due to their own efforts and choices, is an example of analysis conducted at the second level.[71] His claim that justice is the first virtue of social institutions is an example of analysis at the third level, for it gives an account (or a potential account) of how the demands of justice are to be weighed against other values, such as overall welfare.

Rawls is not unusual in operating at all of the three levels of analysis I have distinguished, and there need be no problem in doing so. But one of my criticisms of Rawls has been that he fails to distinguish properly between these different levels and thereby fails to recognize the differences between them and the legitimacy of each. So, for example, he argues that the principle that people should be rewarded in accordance with desert is an inadequate principle of justice since impracticable to implement. But this reasoning is potentially fallacious: it does not follow from the fact that a principle is impracticable that it cannot be a correct principle of justice. At the first level I have distinguished it is perfectly legitimate to maintain that a principle is correct irrespective of whether it can be institutionalized or indeed whether it can be tracked by some set of publicly checkable rules. In maintaining that an adequate theory of justice must be realistic, Rawls also appears to commit himself to a stronger feasibility requirement than can be defended, for at the first level we are entitled to ignore questions about whether, 'under the best foreseeable of conditions', a proposed conception of justice is capable of being the object of a free, reasoned consensus (see earlier).

Identifying these levels of analysis grounds a pluralist approach that allows political theory to operate legitimately at each level, and indeed allows that a particular piece of theoretical analysis may function at all of them. Although there is no reason to establish a division of labour, these levels also permit us to distinguish between political philosophy, understood as the most abstract part of political theory conducted under the fewest empirical constraints, from other parts of political theory which may be more heavily empirically conditioned because they attend to issues of institutional design and involve consideration of the social and political forces that affect the possible implementation of ideals.

(F'note continued)

his constructivist approach seems to identify principles of justice with those principles which are prescribed by a procedure that starts from various facts about human beings and society. It would appear that when Rawls's arguments constitute first level theorizing, they cannot involve the mere application of his constructivist procedure.

[70] See Rawls, *A Theory of Justice*, pp. 72–5/62–5.

[71] See Rawls, *A Theory of Justice*, p. 312/274.

https://doi.org/10.1017/S0007123404000043 Published online by Cambridge University Press