ENCYCLOPEDIA OF GOVERNMENT AND POLITICS

Volume I

Edited by

Mary Hawkesworth

and

Maurice Kogan

ROUTLEDGE
ROUTLEDGE
Taylor & Francis Group

> This gives us the defining characteristic of the second approach, namely, that justice should be the content of an agreement that would be reached by rational people under conditions that do not allow for bargaining power to be translated into advantage.
>
> (Barry 1989:10)

Mutual advantage theory perhaps provides a good analysis of what genuinely rational, purely self-interested people would do. If we are going to engage in amoral *realpolitik* this is perhaps how we should proceed but it does not provide us with anything that even looks like a method of moral justification. A cluster of practices which could be correctly characterized as just practices could not be a set of practices which would sustain or even allow those with greater bargaining power to turn it into such an advantageous outcome that the weak would be killed, die of starvation or live in intolerable conditions of life when that could be avoided. Such practices are paradigmatically unjust practices. If they are not unjust then nothing is.

A mutual advantage theorist might respond that her/his theory could never allow those things to obtain, for, no matter how severe the power differentials, such things (as a matter of fact) would never be to the mutual advantage of the parties (neither the weak nor the strong). But that is clearly a rather chancy empirical claim.[4] Faced, under severe and relatively secure power differentials, with the possibility of starvation, the weak might rationally settle for subsistence wages. Faced with a very marginal subsistence living, families might find it to their advantage (including the children's advantage) to opt for child labour under harsh conditions. With one's back against the wall, one might even find it to one's advantage to sell oneself into slavery or to agree to play a kind of Russian roulette where one might be killed. It is itself a rather chancy empirical claim to say that none of these things would be to the advantage of people in positions of power because the likelihood of the weak sticking with such harshly driven bargains would be too slim. That this would be so in all realistic conditions is far from evident. We can hardly be very confident that positions of power might not be so secure that it would be to the advantage of the powerful to drive such hard bargains. But whatever is in fact the case here about mutual advantage, we can know, impartiality theorists claim, that such bargains are unjust. Thus even if they do turn out to be mutually advantageous, they remain morally unacceptable. To respond 'Well, maybe they won't be mutually advantageous' is not to meet the challenge to mutual advantage theory.

Let us now consider impartiality theories. They take several forms, but whether or not they require the postulation of an original position or a state of nature, such theories view moral reasoning not as a form of bargaining but as a

deliberation or debate between agents who share a commitment to impartiality, to the giving of equal weight to the interests and needs of all. Put differently, they are people who are deliberating about which principles should be acceptable to all points of view. That, as Barry has it, is the basic idea of impartiality. Impartiality theorists such as Rawls, Hare, Sumner, Baier, Nielsen, Barry, Scanlon and Dworkin disagree over which principles of social justice are to be adopted, but they all in some sense are egalitarians and argue (*pace* Hayek) that justice as impartiality requires (where possible) the elimination of morally arbitrary inequalities, namely those inequalities arising from differences in social circumstances or natural talents. How fundamentally such an approach differs from the mutual advantage approach can be seen from the fact that an underlying rationale for appealing to impartial agreement is that it substitutes a moral equality for a physical or intellectual inequality. As Kymlicka well puts it, the two views are, morally speaking, a world apart: 'From the point of view of everyday morality, mutual advantage is an alternative to justice, not an alternative account of justice' (Kymlicka 1990:103).

Appealing here to everyday morality, and not to something more abstract such as the moral point of view, begs the question with mutual advantage theorists, for they are willing to jettison much of everyday morality for a streamlined morality they regard (correctly or incorrectly) as more rational. There are on Hobbesian accounts no natural duties to others, no real moral difference between right and wrong which all persons must respect. There is, as well, no natural moral equality underlying our physical inequality. To the liberal appeal to moral equality (the life of everyone matters and matters equally) the Hobbesian can ask (as James Buchanan does), 'Why care about moral equality?' (Buchanan 1975:54; see also Gauthier 1986).

Hobbesians, to continue the mutual advantage theorist's counter to impartiality theory, will respond to impartialists that they do not push questions of justification to a deep enough level. They do not realize that a person only has a reason to do something if the action the person contemplates doing satisfies some desires of that person, so that if something's being just is to count as a good reason for doing it, justice must be shown to be in the interest of the agent (Barry 1989:363). Keeping this in mind we frame the Hobbesian question of why people possessing unequal power should refrain from using it in their own interests.

To this the impartialist can in turn respond in good Kantian fashion that morality needs no external justification. Morality itself provides a sufficient and original source of determination within us that is no more and no less artificial than the Hobbesian self-interested motivation. People can be motivated to act morally simply by coming to appreciate the moral reasons for doing so.

Hobbesians with their instrumentalist conception of rationality will find this impartialist acceptance here artificial and perhaps evasive. But they in turn must face Barry's claim that to equate rationality with the efficient pursuit of self-interest is a view which rests on pure assertion. Rational egoism is not an inconsistent view. There is no showing that to be consistent one must be an impartialist. But there is no good reason to believe that the very meaning of 'rational' is such that if one is rational one must be an egoist. The acceptance of the formal criterion of universalizability together with a recognition that others are fundamentally like us in having needs and goals and indeed in having, generally speaking, some of the same needs and goals, gives us powerful reasons for accepting the claims of an impartial morality (Barry 1989:273, 285).[5] A person is not being inconsistent if she/he does not care about the needs and goals of others; she/he does not violate the criterion of universalizability, but, as Barry put it, 'the virtually unanimous concurrence of the human race in caring about the defensibility of actions in a way that does not simply appeal to power' (Barry 1989) suggests that this appeal to impartiality and to moral equality are very deeply embedded, considered convictions to some extent held across cultures and over time. To say that such persons act irrationally, if so acting is not in their individual self-interest, or even act in a way that is less than optimally rational, is to utilize what is in effect an arbitrary *persuasive* definition of what it is to be rational.

All constructivist contractualist theories of justice, and of morality more generally, whether mutual advantage theories or impartiality theories, construe justice as those principles and that set of practices on which everyone at least in principle could reach agreement. Barry as much as Rawls construes justice as impartiality in terms of agreement. But there are those who are justice as impartiality theorists but who reject construing justice in terms of agreement (Kymlicka 1989, 1990). Barry gives us a sense of what the stress on agreement would come to:

> [T]he function of justice is to provide a rational basis for agreement among people who do not simply look at things from the point of view of their own interests but seek to take due account of the interests of all. Justice, on this conception, is what can be justified to everyone…. It is inherent in this conception that there is a distinctively moral motive, namely, the desire to behave in accordance with principles that can be defended to oneself and others in an impartial way.
>
> (Barry 1989:272)

Following Scanlon, Barry takes the underlying moral motive to be 'the desire to be able to justify one's actions to others on grounds they could not reasonably reject' (ibid.: 284). Conceptions of this sort are widely held, but Kymlicka among others thinks that they are fundamentally mistaken (Kymlicka 1990:110–12). Perhaps such a conception would work if we were only considering moral

relations between competent adults. But there are as well moral relations between us and children and the mentally disabled. It is senseless to talk about impartial agreement with infants or giving the mentally disabled grounds they could not reasonably reject. Considerations of justice are very stringent between them and us but there is no room for talk of justice coming to what they and we could come to an agreement about.

> If someone is incapable of being a party to an agreement with us, that certainly does not mean we lack any moral motive for attending to his or her interests. The emphasis on agreement within impartiality seems to create some of the same problems that the emphasis on bargaining power creates within mutual advantage theories: some people will fall beyond the pale of morality, including those who are most in need of moral protection.
>
> (Kymlicka 1990:110)

It is a mistake to claim, as Scanlon does, that morality 'only applies to a being if the notion of justification to a being of that kind makes sense' (Scanlon 1982:113–14).

Scanlon maintains in defence of his thesis that the fact that a being can feel pain shows that that being has a centre of consciousness and, because of this, that the notion of justification to such a being makes sense. It is because of this, Scanlon claims, that pain is so often taken as a relevant criterion for moral status. But it is false that if a being can feel pain justification can be addressed to that being and that we can in principle at least attain agreement with her/him. Agreement requires the being not just to be able to feel pain and to be a centre of consciousness, but comprehension as well, and while infants and the severely mentally disabled can feel pain they cannot comprehend things so that they could enter into agreements with us, so the notion of justification would not make sense to them. Yet surely they have moral status. That we cannot address justification to a baby does not mean the baby lacks moral status. We give moral status to an infant not because we can address justification to it or to its moral trustee. We give moral status to it because it can suffer or flourish, because the lives of such beings 'can go better or worse, and because we think their well-being is of intrinsic importance' (Kymlicka 1990:111). Some beings we can address justification to and some we cannot; what 'makes them all moral beings is the fact that they have a good, and their well-being matters intrinsically' (ibid.: 111). But to so argue is to break with the contractarian tradition, including its impartialist versions. But it would seem at least morally arbitrary not to do so.

Kymlicka argues that we should construe justice as impartiality not in the manner of the contractualist as based on some kind of agreement, but that we should simply take impartiality as a criterion that, with or without agreement, gives all interests equal weight. Our moral motivation is not in reaching

agreement but in responding to legitimate interests. We simply come to recognize, if we are moral beings, that others have legitimate claims to have their interests taken into account. The thing is to try to find or articulate principles of justice that give equal weight to everyone's interests. Agreement, Kymlicka claims, drops out.

We have clear obligations to those who are powerless to defend, represent or even recognize their own interests. In this vein, and abstracting a little, our clearest obligations are, Kymlicka claims, not to try to reach agreement but to take people's interests into account and to give equal weight to the interests of all human beings. This is the clear claim of justice as impartiality. Our principles of justice are justified when they do that. If they do not give such equal weight to the interests of all, whether we agree about these principles or not, this agreement does not justify them. This commits us to the substantively egalitarian view that the interests of all human beings matter and matter equally. Where that is not our guiding conception we do not, at least on modern conceptions of justice, have justice. Agreement is, of course, of vital epistemological and political import. But at the foundational level, as Kymlicka has it, it does not apply; that is to say, it does not apply where we are saying what justice is and what the foundations of a just society are (Kymlicka 1990:113). 'At the deepest level,' Kymlicka continues, 'justice is about equal consideration of our legitimate interests, and the many virtues of agreement are assessed by reference to that underlying idea, not vice versa' (ibid.: 112).

There is plainly something right about Kymlicka's argument here, but there may be something wrong as well which gives morals by agreement another inning. What justice as impartiality substantively comes to is giving the interests of all equal weight such that everyone's interests matter and everyone's interests matter equally. Proper names are not relevant in determining whose interest has pride of place when they conflict and both interests cannot be satisfied. Still, in such a situation we must depart from simple equality, and it is there that the careful articulation of principles of social justice such as we find in Rawls, Scanlon and Barry becomes vital. But in making such a differential weighting, such as to proceed by benefiting the worst off maximally in ways that are compatible with retaining autonomy and fair equality of opportunity for all, we should start from a position where we give equal consideration to the interests of all and where we start by giving an initial equal weighting to all interests. It is only when we recognize that not all interests can be satisfied equally that we look for impartial and fair ways of departing from simple equality. But that does not gainsay the point that justice is about the equal consideration of our legitimate interests. This obtains whether or not there is anything that everyone competent to make such judgements and bent on being reasonable would agree on. So far things seem at least to cut against contractarians.

However, let us now ask: how do we know that is so? How do we know that this is what justice is and that this is what justice requires such that we must act in this way if we would be just and that for there to be just social institutions our social practices must be so structured? It is here that agreement may come in by the back door.

Kymlicka writes as if we could just intuit or directly recognize that this is so, that we could just see that these claims are true. But if there are any accounts that are by now widely recognized to be non-starters, it is intuitionism and natural law theories where we in some mysterious way must just have direct access to the truth—indeed, even on some accounts, the certain truth—of certain moral propositions.

How then does Kymlicka know, and how can we know, that his fundamental substantive moral claims, claims not subject to agreement, are true or justified? Perhaps they are (though Kymlicka does not claim that for them) *conceptual* claims such that we can know that they are true by having a grasp of the concept of justice, where to have a grasp of the concept of justice is to know how to use 'justice' or cognate terms correctly. Perhaps the following conceptual chain holds: to be just is to be fair, to be fair is to be impartial, and to be impartial is to give equal consideration to the interests of all human beings. If this is so we could know the truth of Kymlicka's claims by coming to have a good understanding of the use of 'justice'. But that may not give us a way of meeting mutual advantage theories at all. Gauthier, for example, understands perfectly well the ordinary use of 'just' and 'justice' and what it commits us to, if we would stick with it, but he will for his theoretical purposes modify that use until it is compatible with a set of principles that are rationally sustainable and that rational people will agree to be rationally sustainable when these people are reasoning carefully. We cannot go very far in sustaining substantive claims and substantive principles of justice through being clear about the use of 'just' and allied terms. Such considerations may undermine certain absurd claims, but they leave many competitors for what is just in the field.

It may, that is, give us something like the first word but it will not carry us very far beyond that. But then how does Kymlicka know that his substantive claims about justice are justified? He leaves this mysterious. Rawls, Daniels and Nielsen explicitly, and others implicitly, have in such contexts appealed to considered judgements or convictions in wide reflective equilibrium (Rawls 1971:19–21, 48–51, 577–87; 1975; Daniels 1979; Nielsen 1987, 1988b). It has been mistakenly thought that this is a thinly disguised form of intuitionism with all its difficulties plus even more evident worries about ethnocentrism. However, these charges are mistaken, given the kind of coherentism involved in the appeal to considered judgements in wide reflective equilibrium. It starts from our firmest considered convictions of a rather specific sort, such as to enslave people

is wrong, racial prejudice is evil, religious intolerance is unacceptable, and it tries to have a consistent cluster of such beliefs. But it also seeks to show how such specific considered convictions can be derived from and are explainable by more general moral principles, some of which themselves may be considered judgements. 'The interests of all human beings are of equal importance' is one such principle which is also such an abstract considered judgement. We seek by a reciprocal adjusting of many elements, sometimes modifying or abandoning a specific considered judgement or sometimes modifying or even abandoning a more general principle or sometimes by coming to articulate a new one with a powerful rationalizing power, until we get what we can recognize to be a consistent and coherent cluster of beliefs. We do this by sometimes trimming, sometimes expanding, our cluster of considered judgements and principles, but always adjusting this *mélange* of convictions and beliefs. We do this until we have something which we have good reasons to believe forms a consistent and coherent cluster. So far we have nothing more than what is given by ethical intuitionism, though there need be, and indeed should be, no claim to a bizarre epistemic status or a truth capturing power for the moral beliefs and principles. Indeed we can, following Rawls, avoid making any claim about the epistemic status of our principles of justice or our various moral claims.

Where wide reflective equilibrium clearly goes beyond ethical intuitionism, which is a narrow reflective equilibrium, is in its stress that other things besides specific moral beliefs and moral principles must be appealed to in gaining the coherent web of belief and conviction that would constitute a wide reflective equilibrium. The consistent set we seek is not only of specific moral convictions and more general principles, but of whole theories of morality, conceptions of the function of morality in society, factual beliefs about the structure of society and about human nature, beliefs about social change (including beliefs about how societies will develop or can be made to develop) as well as specific historical and sociological beliefs about what our situation is. The equilibrium we seek is one in which all these elements are put into a coherent whole. In narrow reflective equilibrium a specific considered conviction might be abandoned because it conflicted with many equally weighty specific considered convictions or a more general moral principle. But in wide reflective equilibrium they might be rejected as well because they were incompatible with some well-established empirical facts about society or human beings or our particular situation or because they made demands which, given what we know about the world, could not be realized or were beliefs which had moral alternatives which made much more sense in the light of some carefully elaborated social or moral theories or theories about the function of morality in society. There are here a considerable range of considerations, including empirical considerations, that are relevant to our decisions about what to do or how to live. We start with specific considered

convictions but they are correctable by a whole range of empirical and theoretical convictions as well as by moral principles or moral theories, though sometimes in the case of moral principles and theories it will go the other way and the principles or theories will be correctable by the specific considered judgements. This yields a critical morality that lacks the dogmatism and what in effect, though not in intention, is the conventionalism and subjectivism of moral intuitionism. Moreover, that critical morality also functions as a guard against ethnocentrism. Some of the specific judgements we start with may be ethnocentric but by the time we have got them into wide reflective equilibrium the ethnocentrism will be winnowed out.

So if Kymlicka would avail himself of such a procedure he at least arguably would have a method of reason for his fundamental claims of justice and he need not just assert them, somehow taking them to be natural laws or basic intuitions recoverable on reflection. The method of wide reflective equilibrium could, of course, be used, as well, to argue against an account like Kymlicka's. Its advantage, whichever way it is used, is that we do not need to just assert or to rely on intuition with it but can appeal to a method that is very like the method used in science and in other domains.

However, in doing this he would be implicitly appealing to some agreement, to some consensus, for it is *our* considered convictions that we seek to get into wide reflective equilibrium. This means we are in effect appealing to convictions of a specific people, a specific community with its traditions situated in a determinate cultural space and time. We rely on a consensus in such a community though the shared considered convictions need not be, and typically will not be, only the shared convictions–the considered judgements–of that community. They might in some instances be quite pan-human. But for them to be *our* considered judgements they must rest on a consensus in our community and this, of course, implies an agreement. Thus (*pace* Kymlicka) agreement enters in at a very fundamental level. To show that his impartiality account of justice is justified, he must show that its principles and claims can, relying on considered judgements, be placed in wide reflective equilibrium. But this need not mean that it appeals to the agreement of everyone to whom it is addressed.

Some of the philosophers appealing to wide reflective equilibrium, and in doing so relying very fundamentally on considered convictions (Rawls most prominently), are also constructivists and contractarians and take the method of wide reflective equilibrium and their contractarianism to form a coherent whole. For Rawls, for example, in deciding on how thick the veil of ignorance is to be or how the original position is to be characterized, we at crucial junctures rely on considered convictions as we do in deciding on what it is reasonable to accept. But in turn, in deciding on whether we have for a time achieved a reflective equilibrium, we would need a conception of justice which would be acceptable to

the parties under certain idealized conditions. So again, at a very fundamental justificatory level, agreement is appealed to. It is not that the substantive principles and claims of social justice are not what Kymlicka says they are or that justice is what we can agree on in certain idealized situations but that, if we are to show that Kymlicka's or anyone else's substantive claims of justice are justified, we must show that there is such agreement.

We should note in this context that justice is like truth. Truth is not what researchers investigating under ideal conditions and over a considerable time would agree is the case. But that may be the best test for truth. Similarly justice is not what would be agreed to in the original position but that may be the best test for what is just. We have carefully to distinguish what truth and justice mean and what they are from how we ascertain what is true or just.

I want now to consider a way, a rather weak way I am afraid, in which the impartiality approach to justice and the mutual advantage approach *might* be shown to be compatible. The impartiality approaches show us what justice is, how we have to be in order to be just persons of moral principle, what just institutions would look like, and what principles of justice people, reasoning carefully from the moral point of view, would find to be most justified and why. We are asking for moral reasons here which only *per accidens* may sometimes also be reasons of self-interest. Assuming there is something called the moral point of view (one property of which is the impartial consideration of the interests of all), people of moral principle will reason in accordance with it. They will hope and reasonably expect that most of the time their interests will not be hurt by doing so, but they will not think they are justified in doing so only when doing so answers to their own interests or at least does not go against their interests. Their motive for pursuing justice is not the pursuit of individual advantage. What happens to other people matters in and of itself. But we can still ask, and they can ask, *'Why be just?'* Can we give reasons of a broadly prudential sort which will show why a purely self-interested person, if thoroughly rational and clear about the non-moral facts, will do, though out of self-interest, what a just person will do?[6] Kant distinguished between a person of good morals (something an egoist could be) and a morally good person (someone genuinely committed to the moral point of view). Can we show that rational, purely self-interested people, if they were also persons of good morals, would, if they were thoroughly rational, do what just people do, or even do roughly what just people do, though not, of course, for the same reasons? We should recognize in pressing that question that 'Why ought we to be just?', 'Why be fair?', 'Why ought we to do what is right?' or 'Why should we be moral?' are questions that we could not ask from a moral point of view. To ask them is like asking 'Why ought we do what we ought to do?' (Nielsen

However, as the extended discussion of 'Why be moral?' has brought out, we can ask: 'Why take the moral point of view at all?' (Baier 1958; Frankena 1980;

Nielsen 1989c; Gauthier 1988). From the moral point of view, moral reasons by definition override non-moral reasons, but why take that point of view at all? From the point of view of individual self-interest, from class interests or from the point of view of a group of constrained maximizers bent on co-operation for mutual advantage, moral reasons are not the overriding reasons or at most they are only contingently overriding (Wood 1985, 1984). From the moral point of view they are necessarily overriding but not from these points of view. But why take the moral point of view? Justice, fairness and morality requires it. But so what?

Hobbesian theory can be taken as a powerful attempt to show that we have very strong prudential reasons for being, as the world is and will continue to be, persons of good morals. We have in terms of long-term self-interest the best of reasons to support the continued existence of moral constraints, including just practices. (We could not–logically could not–have moral institutions, at least where the circumstances of justice obtain, that did not include just social practices.) Rational persons, the claim goes, will not be morally good, but they will be of good morals.

The impartialist arguments, such as we have seen Barry and Kymlicka articulating, show, I believe, that Hobbesians (pure mutual advantage theorists) cannot get justice out of purely self-interested reasoning, including constrained maximization, which in the end is itself purely self-interested reasoning.[7] Indeed, it is true, as some modern Hobbesians have powerfully argued, that people can expect to advance their interests most effectively by co-operating and in doing this by agreeing to accept certain constraints on their direct individual utility maximization. By moderating their demands and by cooperating with others they will, as the world goes, in the long run do better. David Gauthier makes a powerful case for that (Gauthier 1986).[8] But these forms of co-operation will not give us morality, will not give us a system of justice, where the interests of all count equally, where what happens to other people matters in and of itself, where the reasons for action must not just be acceptable from the point of view of the agent doing the reasoning but from all points of view. For a social practice to be just it must not only answer to the interests of some individual or some class or elite but it must also answer to the interests of all. But, as we have seen, there can be all kinds of situations (class differentials, caste systems, hierarchical strata, adults and children, the mentally competent and the mentally disabled, developed cultures and non-literate ones) where there are differential power structures and where, by pursuing mutual advantage intelligently in certain circumstances, the powerful would exploit the weak and not for all of that be acting unintelligently. It could, as we have seen, very well in such circumstances be in the mutual advantage of everyone involved. Justice cannot allow differential bargaining power to be translated into advantage: that is, it cannot allow exploitation. People in such circumstances, given their weakness, have reason to

co-operate with the strong for otherwise they will be still worse off. And in societies as we know them these circumstances are not infrequent. So, given the differential power situation and the determination of the powerful to do the best they can for themselves, the weak have prudential reasons to co-operate even though they are exploited. But they are not being treated justly; the resulting system of co-operation, though rational, is not moral. Indeed such treatment of people is immoral. We do not reach morality from Hobbesian premises and thus we do not reach justice. The impartialist does not ask why be just but shows what justice is; the Hobbesian asks why be just and tries to show that we should be just because justice pays. What has been shown is that it is not true that justice *always* pays. Some form of social co-operation always pays, but the form of social co-operation people engage in may be very different fran justice. The Habermasian has not shown that the enlightened egoist or the intelligent and informed constrained maximizer must, to be thoroughly rational, be just. But the Hobbesian has not shown that we can get justice out of enlightened egoism.

To this the Hobbesian might reply that a good bit of morality is irrational. The moral point of view requires the equal consideration of interests but it is irrational for an individual or a group to do so when it is not in their interests. What is rational to do is determined by the interests of the individual who is doing the acting. Where parts of morality do not so answer to individual interests they should, the Hobbesian can claim, be jettisoned and what is kept as a system of social co-operation, though considerably less than morality as it has been traditionally conceived, is the rational critical core of morality.

This purely instrumentalist conception of rationality, as we saw Barry arguing, is pure assertion. That it is just this that rationality comes to is not established through an examination of the use of 'rationality'. To give equal weight to the interests of all is not irrational. To say it is a rational thing to do is no more or no less rooted in the use of 'rational' than is the claim that to be rational is always to give self-interested reasons pride of place.

We can appeal to theoretical considerations to support such an instrumentalist conception of rationality, but there are other conceptions of rationality answering to different theoretical purposes. Given Hobbesian purposes we can use that Hobbesian conception of rationality, but, given Habermasian or Aristotelian purposes or the purposes of impartialism, we can use instead these quite different conceptions of rationality. There seem to be no good reasons external to these particular purposes to accept one of these purposes rather than another; and to say that the Hobbesian ones are the really rational ones is plainly question begging. Moreover, the Hobbesian conception is subject to *reductio* arguments. *If* it fits the interests of one class to enslave another class and work them to the edge of starvation, that would, on such a Hobbesian account, not only be what reason *permits;* it would be what reason *requires,* but a theory of rationality that had that

implication would not only be morally repugnant, it would be groundless and thoroughly implausible.[9]

NOTES

1 Nozick (as do many other libertarians) takes himself to be a genuine descendant of Locke. This has been impressively challenged by Virginia Held (1976) and Shadia Drury (1979).

2 Gauthier's position is the canonical one here. Narveson's far cruder politically committed work seeks to follow Gauthier. It is a question worth pursuing to ask how much, if any, of Herzog's trenchant critique of Narveson rubs off on Gauthier.

3 The latter claim is Habermas's and, unlike Rawls, he is not loath to make meta-ethical claims. However Rawls, with his method of avoidance, does resolutely set aside meta-ethical claims.

4 This is reminiscent of utilitarian arguments to ward off *reductio* arguments against utilitarianism.

5 To say that something is universalizable is to say that, if X is good for Y or is something Y ought to do, it is something that is good for anyone else or something anyone else ought to do if that someone is relevantly like Y and is relevantly similarly situated. 'Relevantly' here needs to be cashed in contextually. See Nielsen (1989b).

6 I am inclined to think 'non-moral facts' is pleonastic but that belief is contentious.

7 Gauthier remarks 'my discussion assumes rational, utility-maximizing individuals who are not mistaken about the nature of morality or, more generally, who recognize that *the sole rationale* for constraint must ultimately be a utility-maximizing one' (Gauthier 1988:182).

8 The work of Will Kymlicka has deeply influenced me in the writing of this essay. His influence, my criticisms to him to the contrary notwithstanding, is particularly evident in the last third of this essay.

9 This contention about rationality is elaborated and defended in Nielsen (1991).

BIBLIOGRAPHY

Aristotle (1975) *The Nicomachean Ethics,* Cambridge, Mass.: Harvard University Press.

——(1977) *Politics,* Cambridge, Mass.: Harvard University Press.

Baier, K. (1958) *The Moral Point of View,* Ithaca, NY: Cornell University Press.

Barber, B. (1988) *The Conquest of Politics,* Princeton: Princeton University Press.

Barry, B. (1989) *Theories of Justice,* Berkeley: University of California Press.

Bell, D. (1979) 'On meritocracy and justice' in D.L.Schaefer (ed.) *The New Egalitarianism,* Port Washington, NY: Kennikat Press.

Buchanan, J. (1975) *The Limits of Liberty: Between Anarchy and Leviathan,* Chicago: University of Chicago Press.

Cohen, G.A. (1989a) 'Are freedom and equality compatible?', in J.Elster and K.O. Moene (eds.) *Alternatives to Capitalism,* Cambridge: Cambridge University Press.

——(1989b) 'On the currency of egalitarian justice', *Ethics* 99 (4):906–44.

Daniels, N. (1979) 'Wide reflective equilibrium and theory acceptance in ethics', *Journal of Philosophy* 76:256–82.

Drury, S. (1979) 'Robert Nozick and the right to property', in A.Parel and T.Flanagan (eds) *Theories of Property,* Waterloo, Ontario: Wilfred Laurier University Press.

Dworkin, R. (1985) *A Matter of Principle,* Cambridge, Mass.: Harvard University Press.

Elster, J. (1984) 'Exploitation, freedom and justice', in J.R.Pennock (ed.) *Marxism, Nomos* vol. 26, New York: New York University Press.

Frankel, C. (1971) 'Equality of opportunity', *Ethics* 81 (3):191–211.

Frankena, W. (1980) *Thinking about Morality,* Ann Arbor: University of Michigan Press.

Gauthier, D. (ed.) (1970) *Morality and Rational Self-Interest,* Englewood Cliffs, NJ: Prentice-Hall.

——(1985) 'Justice as social choice', in D.Copp and D.Zimmerman (eds) *Morality, Reason and Truth,* Totowa, NJ: Rowman & Allenheld.

——(1986) *Morals by Agreement,* Oxford: Clarendon Press.

——(1988) 'Morality, rational choice and semantic representation', *Social Philosophy and Policy* 5 (2):182.

Gilbert, A. (1990) *Democratic Individuality,* Cambridge: Cambridge University Press.

Habermas, J. (1983) *Moralbewußtsein und kommunikatives Handeln,* Frankfurt am Main: Suhrkamp Verlag; pub. in English (1990) as *Moral Consciousness and Communicative Action,* trans. C.Lenhardt *et al.* Cambridge, Mass.: MIT Press.

Hampshire, S. (1989) *Innocence and Experience,* Cambridge, Mass.: Harvard University Press.

Hardin, R. (1988) *Morality within the Limits of Reason,* Chicago: University of Chicago Press.

Hare, R.M. (1978) 'Justice and equality', in J.Arthur and W.Shaw (eds) *Justice and Economic Distribution,* Englewood Cliffs, NJ: Prentice-Hall.

Hayek, F. (1960) *The Constitution of Liberty,* London: Routledge; Chicago: University of Chicago Press.

——(1973) *Law, Legislation and Liberty,* 3 vols, Chicago: University of Chicago Press.

Held, V. (1976) 'John Locke on Robert Nozick', *Social Research* 43:169–95.

Heller, A. (1987) *Beyond Justice,* Oxford: Basil Blackwell.

Herzog, D. (1990) 'Gimme that old-time religion', *Critical Theory* 4 (1, 2):74–85.

Kain, P.J. (1988) *Marx and Ethics,* Oxford: Clarendon Press.

Kymlicka, W. (1989) *Liberalism, Community and Culture,* Oxford: Clarendon Press.

——(1990) 'Two theories of justice', *Inquiry* 33 (1):81–98.

Levine, A. (1984) *Arguing for Socialism,* London: Routledge & Kegan Paul.

Locke, J. (1970) *Two Treatises of Government,* ed. P.Laslett, Cambridge: Cambridge University Press.

MacIntyre, A. (1988) *Whose Justice? Which Rationality?,* Notre Dame: University of Notre Dame Press.

Marx, K. (1962) *Critique of the Gotha Programme,* ed. C.P.Dutt, New York: International Publishers.

Miller, R. (1981) 'Marx and Aristotle: a kind of consequentialism', in K.Nielsen and S.Patten (eds) *Marx and Morality,* Guelph: Canadian Association for Publishing in Philosophy.

Nagel, T. (1979) *Mortal Questions,* New York: Cambridge University Press.

Narveson, J. (1988) *The Libertarian Idea,* Philadelphia: Temple University Press.

Nielsen, K. (1985) *Equality and Liberty: A Defense of Radical Egalitarianism,* Totowa, NJ: Rowman & Allenheld.

——(1987) 'Searching for an emancipatory perspective: wide reflective equilibrium and the hermeneutical circle', in E.Simpson (ed.) *Anti-foundationalism and Practical Reasoning,* Edmonton: Academic Printing and Publishing.

——(1988a) 'Radically egalitarian justice', in A.F.Bayefsky (ed.) *Legal Theory Meets Legal Practice,* Edmonton: Academic Printing and Publishing.

——(1988b) 'In defense of wide reflective equilibrium', in D.Odegard (ed.) *Ethics and Justification,* Edmonton: Academic Printing and Publishing.

——(1989a) *Marxism and the Moral Point of View,* Boulder, Colo.: Westview Press.

——(1989b) 'Justice, equality and needs', *Dalhousie Review* 69 (2):211–27.

——(1989c) *Why be Moral?,* Buffalo, NY: Prometheus Books.

——(1991) 'Can there be justified philosophical beliefs?', *Iyyun* 12:69–94.

Nozick, R. (1974) *Anarchy, State and Utopia,* New York: Basic Books.

Okin, S.M. (1989) *Justice, Gender and the Family,* New York: Basic Books.

Peffer, R. (1990) *Marxism, Morality and Social Justice,* Princeton: Princeton University Press.

Pogge, T.W. (1989) *Realizing Rawls,* Ithaca, NY: Cornell University Press.

Rawls, J. (1971) *A Theory of Justice,* Cambridge, Mass.: Harvard University Press.

——(1975) 'The independence of moral theory', *Proceedings and Addresses of the American Philosophical Association* 47:7–10.

——(1978) 'The basic structure of the subject', in A.Goldman and J.Kim (eds) *Values and Morals,* Dordrecht: Reidel.

——(1980) 'Kantian constructivism in moral theory', *Journal of Philosophy* 77:515–72.

——(1985) 'Justice as fairness: political not metaphysical', *Philosophy and Public Affairs* 14 (3):223–51.

Reiman, J. (1990) *Justice and Modern Moral Philosophy,* New Haven: Yale University Press.

Riley, P. (1982) *Will and Political Legitimacy: A Critical Exposition of Social Contract Theories in Hobbes, Locke, Rousseau and Hegel,* Cambridge, Mass.: Harvard University Press.

Roemer, J. (1988) *Free to Lose,* Cambridge, Mass.: Harvard University Press.

Sandel, M. (1982) *Liberalism and the Limits of Justice,* Cambridge: Cambridge University Press.

Scanlon, T. (1982) 'Contractualism and utilitarianism', in A.Sen and B.Williams (eds) *Utilitarianism and Beyond,* Cambridge: Cambridge University Press.

Schofield, M. and Sorabji, R. (eds.) (1977) *Articles on Aristotle: 3–Ethics and Politics,* London: Duckworth.

Shklar, J. (1986) 'Injustice, injury and inequality', in F.Lucash (ed.) *Justice and Equality Here and Now,* Ithaca, NY: Cornell University Press.

Sumner, W. (1987) *The Moral Foundation of Rights,* Oxford: Oxford University Press.

Taylor, C. (1985) *Philosophical Papers,* vol. II, Cambridge: Cambridge University Press.

Tully, J. (1982) *A Discourse on Property: John Locke and his Adversaries,* Cambridge: Cambridge University Press.

Walzer, M. (1983) *Spheres of Justice,* Oxford: Basil Blackwell.

Wood, A.W. (1984) 'Justice and class interests', *Philosophica* 3 (1):9–32.

——(1985) 'Marx's immoralism', in B.Chavance (ed.) *Marx en Perspective,* Paris: Éditions de l'École des Haute Études en Sciences Sociales.

6

CONCEPTIONS OF HUMAN NATURE

LESLIE STEVENSON

Theories of human nature attempt to identify and explain the fundamental features of the human species; and many theorists go on to offer prescriptions as to how human life ought to be conducted, both at the level of individual behaviour, and the level of social and political policy. There has been intense disagreement about a number of basic issues: whether humans are essentially different from other animals; whether they differ importantly from each other (individually, or in races or other groups); whether human nature is constant, or historically and culturally variable; whether human nature is basically good and in need only of appropriate sustenance, or in important respects defective and requiring transformation. There has, as a result, been much argument about the role of government and politics in sustaining or changing human life.

The multiple ambiguity of the term 'nature', as used in this whole debate, should be noted straightaway. In asking how far human nature can be changed, we usually mean human dispositions and behaviour as we know them, in the society we presently live in. But some influential thinkers—notably Hobbes, Locke and Rousseau—have used the phrase 'human nature' (or its equivalents) to express their conception of how human beings would behave if there were no society, no state, government or politics, and presumably little or no education or culture. Sometimes the conception is expressed historically, in a claim about how things were before the beginning of government. The contrast has been variously expressed as between the given and the artificial, the natural and the conventional, the biological and the social, the original and the present day.

Another important ambiguity is about whether the supposed natural state of humanity is to be preferred or avoided. In contemporary discourse what is 'natural' is often assumed to be good (as in natural yoghurt, natural colours, natural lifestyles); certainly what is described as 'unnatural' is thereby condemned as bad. Hobbes famously presented the pre-social 'state of nature' as

'nasty, brutish and short', and saw the social contract as the only rational way of escape from it. Both he and Locke use the state of nature as a device to illuminate the advantages of political society, and to justify certain relationships of authority. But Rousseau, writing about a century later (against the prevailing optimistic mood of the Enlightenment), argued that society had introduced all sorts of unjust inequalities. In his early work the state of nature serves as a critique of many of the crucial features of existing society, and it is easy to see how (in the era of the French revolution) his conceptions could be used to support attempts at radical reform. Rousseau has probably been influential in fostering the idea that what is 'natural' must therefore be best, but it is a highly contentious assumption.

This essay will provide a brief overview of some of the most politically influential conceptions of human nature, noting how normative views can be concealed within apparently factual theories, and comparing them on the issue of constancy versus changeability. Some theorists have held that human nature *could* be substantially altered, given sufficiently radical changes in political or economic structures, or in social practices such as infant-rearing, education, or religious observance. We can call those who offer such remedies 'social engineers', in that they hold that human behaviour could be substantially changed for the better, and human beings made happier, if only their recommended social set-up could be instituted. But other theories, whether biological, social or theological, imply that there are strict limits to how far human nature can be affected by variations in social conditions. The debate here has wide ramifications—into political and social theory, sociology, psychology, biology, philosophy and theology. It is not, however, a lining-up behind simple 'yes' or 'no' answers as to whether human nature can be changed, for we cannot do justice to the different views by trying to divide them neatly into 'constantists' and 'variabilists'. There is, rather, a great variety of views about how far, and under what conditions, human nature might be changed, and how much it must remain the same. So we may as well review our selected theories in historical order.

PLATO

More than two thousand years ago, the Greek philosopher Plato set out a very influential description of an ideal society in his lengthy dialogue, *The Republic*. His discussion ranges very widely, from psychology, metaphysics and moral philosophy to education, art and the status of women. Plato's theory of individual human nature is that in each person there are three mental factors at work—Reason (rationality), Appetite (bodily desires), and Spirit (which is something like courage, pride or personality). These elements each have their proper part to play, but they can sometimes conflict, and what is needed for

human flourishing is a harmonious combination of them, with Reason in firm overall control. Different people will have different factors more strongly represented, so there is no natural equality between individuals.

Neither does Plato think there should be social or political equality—thus opposing the democratic tendency of the Athens of his time. For he argues that the best way for society to be organized is for those with the most developed Reason to have authority and power, since they know what is best—it should not be a matter for mere counting of opinions or preferences amongst everyone. In fact, he proposes a strict threefold class division in society, affecting lifelong duties and status, paralleling his tripartite theory of the human mind or soul. There is to be a class of Rulers or Guardians (carefully selected and trained), a class of Auxiliaries which comprises all state-functionaries including soldiers, police and civil servants, and a class of Workers in all trades, agricultural or urban. Plato thinks that society can be stable and harmonious only if each class of people is restricted to their own special function. The trained elite has a duty to rule, even if they would prefer to spend their time in philosophical thought (and they are not to be permitted either families or private property), whereas the Auxiliaries and Workers have no business in ruling, not even in voting for prospective rulers, for they lack all relevant knowledge. For Plato the well-being of the society does not consist in the well-being of its individual members. There is a certain totalitarian air about his ideal republic, revealed also in his recommendations of strict censorship of the arts, to prevent any destabilizing ideas gaining currency.

An elaborate, deeply argued philosophical vision—the theory of Forms as perfect, eternal, unchanging objects of knowledge grasped by the Reasoning element within the human soul—lies behind Plato's conception of knowledge. He implies that what we would now call questions of value, about what is best for individuals and for societies, can be as much matters of knowledge as propositions in mathematics or science. The obvious difficulty for this idea is the widespread, and apparently irresolvable, disagreement that exists (then, and now) about most questions of value. If there are facts about such matters, facts which are knowable by human beings, why the persistent disputes? Plato realizes that there is considerable difficulty in attaining the relevant 'expertise', and he prescribes a detailed programme of education (restricted to those capable of benefiting from it) by which the future Guardians, the 'philosopher-kings', are to be trained. But he can offer no guarantee that even the best-educated elite will always govern in the interest of society as a whole, rather than in their own interest, and he offers no mechanism for changing rulers, or for resolving disputes between them.

Plato's conception is thus a remarkably unpolitical one. He did not say how in real-life politics his prescriptions can be put into practice or be maintained—it is as if he hoped that their intrinsic rationality would persuade people to accept them.

(His attempts to apply his theory, when given the chance to educate prospective rulers of Sicily, were notoriously unsuccessful.) His is a timeless, transcendent, other-worldly kind of theory, with no allowance for human dispositions such as family ties which do not fit into his ideal state, no provision for failures to fulfil the social functions he allots, and no recognition of variations between people and societies at different times or places.

HOBBES

Writing about the time of the English civil war in the mid-seventeenth century, Hobbes, in his *Leviathan,* presents pre-social human life as extremely insecure, because of the constant danger of fighting over vital resources. He bases his description of individual human nature on a strictly materialist conception–which he thinks is required by the new methods of physical science–of humans as consisting of nothing more than matter in motion. In Hobbes's view, each individual is purely self-interested, seeking the satisfaction of his or her present desires, and the acquisition of means for future satisfaction: 'I put for a general inclination of all mankind, a perpetual and restless desire of power after power, that ceaseth only in death.' There is no co-operation (except when it serves individuals' self-interests), just a constant competition between individuals of approximately equal strength and intelligence. So even when in possession of house, crops, animals, etc., there will always be fear that these will be forcibly taken by someone else; and this gives each person reason to make pre-emptive strikes against others, extending power in order to increase security. People even come to value power over others for its own sake, and to enjoy 'reputation' (Hobbes shrewdly observes that reputation of power *is* power, since it influences how people act). So without any 'common power to keep them all in awe', people live in a state of war with every individual against everyone else, not always actually fighting, but in constant fear of it. In this condition there is little incentive for any longer-term projects like agriculture, industry or science. There can be no applicable notions of justice, rights, property or law; there is only the fact of physical possession until dispossessed by superior force.

Agreements between individuals are of no use in remedying the state of nature, for when it is in someone's self-interest to break such an agreement, what reason do they have to keep it? 'Covenants, without the swords, are but words, and of no strength to secure a man at all.' In Hobbes's view, this gives each person an overwhelming good reason to accept a social contract by which all subject themselves to the supreme power and authority of a 'sovereign'. 'The only way to erect such a common power, as may be able to defend them from the invasion of foreigners, and the injuries of one another…is to confer all their power and strength upon one man, or one assembly of men.' Thus is created a

'commonwealth' (the 'leviathan' of the title of Hobbes's famous book), or what we would now call a state, with a government. Note that this need not be thought of as a historical event: the main point is to show why everyone has good reason to accept the authority of the state (provided that there *is* a single source of power that is effectively unchallenged). The implication of the argument is that any state authority is better than none, and that those that are in actual control deserve allegiance because of that fact alone.

Hobbes's account of the authority he thinks the sovereign (or sovereign body) must have is remarkably authoritarian. Those who are subjects of a monarch have no rights, without his permission, to 'cast off monarchy', that is, to cancel the contract and become a member of another state or of none. And because the contract is between individuals themselves, not between individuals and sovereign, Hobbes says there can be no such thing as breach of contract by the sovereign; he may commit 'iniquity', but not 'injustice'. Further, the sovereign has the right to judge which opinions are dangerous to the state, and may censor publication of them. The sovereign is to make laws and administer them; to conduct foreign policy and decide on war and peace; to appoint all government officials; and distribute reward or punishment as he or she pleases. Hobbes makes no provision against misuses of power: he seems so afraid of the horror of the 'state of nature', as he sees it, that he is prepared to risk despotism to avoid it.

LOCKE

Just a few decades later—about the time of the 'Glorious Revolution' of 1688 in England, by which the power of the monarchy was limited—Locke, in his second *Treatise on Government,* paints a less dark picture of the 'state of nature', and presents the introduction of government more as a matter of convenience than dire necessity. To an extent, he admits (like Aristotle) that human nature is already social, that we are so made that 'it is not good for us to be alone', being naturally disposed to live not merely in families, but as members of wider groupings. However, he still uses very freely the idiom of a pre-social, or at least pre-governmental, state of nature.

Locke conceives of people in this condition as being both free and equal, in that nobody has more power or authority than any other, but he differs from Hobbes in holding that this can be a state of 'peace, goodwill, mutual assistance and preservation'. Another difference from Hobbes's state of nature is that Locke posits a fundamental notion of property, with the distinctive rights of use and disposal, as a corollary of human existence, even in the pre-social state. Whatever someone 'mixes his labour with' for personal use, for example plucking a wild fruit, cultivating crops, or digging ore from the ground, becomes private property: 'as much as anyone can make use of to any advantage of life

before it spoils, so much he may by his labour fix a property in'. Clearly, Locke is optimistically assuming that there is no scarcity of vital necessities in the 'state of nature' (he refers to the contemporaneous settlement of almost uninhabited regions of America). Hobbesian competition for resources is surely probable as soon as human population outstrips the capacity of the environment to sustain it, but Locke can claim that human beings are not *inevitably* aggressive towards each other, and that in conditions of economic sufficiency they will not be. According to Locke, there is a 'law of nature' which applies even in this pre-social condition, since rational beings are able to realize that 'no one ought to harm another in his life, health, liberty, or possessions' (he tries to back this up with a pious appeal to the wise intentions of the divine Creator). But he is not so naïve as to suppose that everyone will readily obey this law, and so he maintains that in the pre-social state everyone has a right to punish transgressions of the law of nature, and the injured party has a particular right to take reparations from the offender.

This is the point in Locke's argument where government comes in. Recognizing that it is dangerous to let individuals be judges in their own cases, since they will easily be led into punishment beyond what is justified, he says that civil government is 'a proper remedy for the inconveniences of the state of Nature'. But having learnt from his experience of the Stuart kings, he notes that absolute monarchs can abuse their power. And, in a crucial criticism of Hobbes, he argues that far from being a remedy for the state of nature, absolute sovereignty is no escape from it at all, since individual and sovereign are really in a state of nature with respect to each other as long as there is no legal check on the power of the latter over the former. Locke is thus a foremost theorist of how the legitimacy of government must depend on the consent of the governed, and of how all power needs to be subject to restraint; his ideas strongly influenced the Constitution of the United States of America.

Hobbes and Locke differ in their conceptions of pre-social human nature, and so (it seems) they diverge over what political arrangements they recommend. Or is it really the other way round—that because they have different political views (Hobbes favouring absolute authority, and Locke wanting checks on state power), they think up different theories of human nature to try to justify these views? There is no serious attempt by these writers to find out the facts about the pre-history of humankind, or about how people would behave if there were no state power. It looks very much as if what are presented as factual, even scientific, descriptions of human nature already conceal within themselves the normative preconceptions of their authors—a possibility to which we must be alive in other theories.

ROUSSEAU

In his *Discourse on Inequality,* Rousseau seems to make more of an effort than Hobbes or Locke to paint a historically realistic picture of the stages by which present society must have evolved from the primeval human condition. He refers to some of the zoological reports of exotic creatures and anthropological evidence about primitive cultures which were then circulating in Europe. He speculates about how human language might have evolved out of instinctual cries. He accuses Hobbes of reading back into the state of nature motives like pride which can only exist in society, and he claims (also against Hobbes) that humans have an innate repugnance against seeing a fellow creature suffer, which moderates the competition between individuals. Rousseau's description of 'the noble savage' represents humans as 'wandering in the forests, without work, without speech, without a home, without war, and without relationships', and this 'without any need of his fellow men and without any desire to hurt them'. There was no inequality between individuals, except relatively small differences in strength, intelligence, etc. There was neither education nor historical progress; each generation lived as its ancestors had done.

Rousseau goes on to speculate about our evolution since then. He treats the notion of property, rather than political power, as most distinctive of civil society. He suggests that the true golden age was at the stage when people had come to form families living together in houses, with some degree of interfamilial socialization into communities, property rights recognized for the immediate necessities of life, and offences against these punished–very much Locke's state of nature, in fact. This for Rousseau was 'the true youth of the world', and he interprets all so-called progress since then as really steps towards 'the decrepitude of the species'. He blames the division of labour, especially in agriculture and metallurgy, for starting the rot, making it necessary for many people to work under the direction of others, allowing some to amass huge property, and thus making possible all the manifold forms of exploitation and economic and social inequality of which he was so painfully conscious. His analysis in this work is a tragic one–that the economic progress due to human cleverness has also developed wickedness, and brought out the worst in human nature. But one suspects that his revulsion from certain features of the society he knew leads him to idealize his speculative 'golden age'.

In that work Rousseau did not offer much in the way of a recommendation for how to cure or alleviate the unhappy condition which he diagnosed in society, there being no realistic possibility of a return to the past. But in his later work, especially *The Social Contract,* he took a more positive view, arguing that human nature does after all find its most complete fulfilment in civil society, at least at its best. Like Hobbes and Locke, Rousseau uses the device of a 'social contract' to

explain the allegiance owed to political authority. People in the state of nature are supposed to reach a critical stage where they realize that their very survival is at risk, and to find it each to their advantage to enter into an agreement with everyone else. But in Rousseau's version the power is granted not to a Hobbesian absolute sovereign, nor even to an elected government, but rather to the community as a whole, which becomes a moral entity in itself. And this involves his distinctive, but rather mysterious notion of the 'general will', which is always for the good of the whole, and yet cannot be identified with the actually expressed will of the citizens, even if all of them should vote in an assembly. But at this point a theory of human nature as it is ceases to play a role in Rousseau's thought: the 'general will' has to be what people *ought* to want, not what they actually want. Such a notion makes it all too easy for those in power to claim that they know better than the people what would be good for them.

MARX

Karl Marx, writing in the nineteenth century when ideas of historical evolution were all the rage, presents a wide-sweeping theory of the development of human societies through various stages, characterized primarily by the nature of their economic production—from the ancient cultures, through the feudalism of the Middle Ages, into the capitalist mode of production, to be superseded (he predicted) by a revolutionary change to the communist mode. According to Marx's conception of human nature, humans are essentially social beings, who do not merely find their means of subsistence in the world but have to work to produce them—for example, growing crops, domesticating other animals, building shelters and making tools. From this emerges Marx's claim that the specific characteristics of a determinate population depend on the kind of society they are members of, which depends in turn on the existing mode of production of the necessities of life.

Marx presents this 'materialist theory of history' as an objective, scientific analysis of the laws governing human societies. He was not, however, merely a dispassionate academic theoretician, he was keenly aware of what he saw as the grave injustices of the capitalist society of his day. He not only predicted, but longed for, the transition to communism, in which he believed that a system of common ownership of the means of production would allow, for the first time in history, the free development of the potential of all human beings. Although, according to Marxist theory, the revolution could not happen until the economic development of a society made the time ripe, as that time approached there would be opportunity for those with an accurate understanding of the situation to prepare the way by organization and propaganda, and when the chance came, to seize the initiative and bring about the revolutionary transfer of power to the

communist party, as Lenin did in 1917. It is only in this sense that Marx can be said to be a social engineer. As to how things were to proceed after the revolution, he was optimistic but very vague; he foresaw the need for a 'dictatorship of the proletariat' for a transitional period, but after then he thought that the state could 'wither away'.

Experience (at least until very recently) has shown quite the reverse happening: the dictatorship of the communist party (the self-appointed representatives of the proletariat) strengthened to totalitarian terror, social engineering was undertaken on a huge scale, and state power extended into almost every feature of life. The Marxist analysis of human nature tends to ignore the persistence of certain kinds of human behaviour even through fundamental economic and political changes—the enjoyment of power and privilege by individuals and ruling groups, the rivalries engendered by nationalist and ethnic feeling, and the desire of many to engage in economic enterprise for themselves.

SOCIAL DARWINISM

In stark contrast to the Marxian conception of human nature, 'Social Darwinism' (which underlies the pronouncements of the more ideological defenders of the 'free market economy') offers an account that enshrines competition as both inescapable and desirable in human life. Darwin himself cannot be held responsible for this view—his theory of evolution by natural selection is an explanation of the origin of the diversity of all living species, not itself a theory of human society. However, since the time of Herbert Spencer in England and W.G.Sumner in the USA (see Jones 1980, Rose 1984), political and social theorists who favour the least possible control by the state over economic activity (the doctrine of *'laissez faire'*) have often appealed to certain Darwinian ideas to try to justify their prescriptions. (They can count as social engineers only in the Pickwickian sense that in countries where there has been a tradition of state-managed economy and social services they will want to change these institutions, and so this programme can constitute a revolution of sorts.)

Their creed can be seen as implicit in the phrase 'survival of the fittest' (the words are Spencer's, not Darwin's). This is to be read not merely in the factual, Darwinian sense that only those individuals best fitted to the prevailing environment will survive (or at least, live long enough to leave progeny), but in the normative sense that it is a good thing that this should be so, and that the less fit should not survive, or not survive so well or so long. It is a political ethic that makes a virtue out of competition; and it obviously suits the successful capitalist very well, for it seems to justify ruthless elimination of rivals, to bless economic

success with virtue on top of material reward, and to discourage any attempt at redistribution of resources through taxation or any other compulsory measure.

But it does not amount to much of a theory of human nature, for all it does is to point to competitive tendencies in economic activity as one aspect of human behaviour, to claim that these can work to the benefit of everyone, and then to jump to the sweeping conclusion that individual economic freedom is the only thing that is important. It leaves out of consideration all co-operation between people, indeed it seems to treat individuals or families as isolated units without acknowledging membership in larger social groups which have profound effects on individuals' identities, obligations and rights.

SKINNERIAN BEHAVIOURISM

A conception of human nature supporting large-scale social engineering has been extrapolated from the behaviourist psychology of the American psychologist B.F.Skinner, whose theories have had some limited success in explaining and modifying the behaviour of various species of animal under laboratory conditions. In this case the claim for applicability to the problems of human society has been made by Skinner himself, but just what he proposes remains rather vague (Skinner 1953). He believes inherited factors play a fairly small role in determining behaviour, and like Marx he strongly emphasizes the plasticity of human behaviour to social influences (which Skinner will label 'conditioning'). But unlike Marx he suggests that regardless of the historical and economic background, knowledgeable behavioural scientists can intervene to create whatever kind of people are wanted, simply by arranging the conditioning influences accordingly. He thus proposes that social scientists 'design a culture' to optimize individual and social benefits, dispensing with troublesome notions of individual freedom and responsibility as 'unscientific'. On this view, human beings are merely creatures whose behaviour is determined by conditioning influences from their past and present social environment.

Clearly, this leaves very much open just what sort of people and society we should be trying to create; on this point Skinner is much less explicit than Plato, and his view seems to amount to no more than the offer of a behavioural technology (which many would argue has little application to human beings, since there are species-specific limits to conditioning) towards ends or goals which remain unspecified—and could in practice turn out to be those of the commercial advertiser, the religious evangelist, the ruling party's propagandists, or whoever else is able to get access to the main means of conditioning people (such as television).

SOCIO-BIOLOGY

Let us turn from the modern social engineers, the optimists about the transformation of human behaviour through social change, to those who emphasize the fixity of human nature. Prominent recently have been those who take a firmly biological view of human beings as one species amongst others, and claim that the important determinants of our behaviour are innate, bred into us by our evolution and coded in the molecules of our genes.

Let us briefly mention Freud as an interesting intermediate case here. He was a pioneer of the biological approach to human nature, putting forward a theory of instincts, while also emphasizing the importance for character-formation of the early years of strong attachment to parents. He claims to detect the unconscious, instinctual influences behind human behaviour, often dismissing as mere 'rationalizations' the reasons explicity offered. But in practical therapy, Freud appears as more of a rationalist—the aim of his distinctive 'psycho-analytic' treatment being to bring into consciousness, for free rational decision, the features which had been repressed into the unconscious mind. Freud sometimes suggested the applicability of his theories to social questions. But nothing in the way of a social programme or political creed can be ascribed to him, only the general thought that there has to be a compromise between society and individuals. Civilization requires the giving up of some instinctual satisfaction—but if it is to exist at all, allowance must be made in our social arrangements for the innate, unchangeable nature of humankind.

This biological theme has been taken up by others who have studied human beings as one kind of animal amongst others—ethologists such as Konrad Lorenz and, more recently, self-styled 'socio-biologists' such as Edward O. Wilson. Lorenz offers a controversial diagnosis of human aggressive tendencies, on the basis of his theory of intraspecific aggression in a variety of animal species. He explains it as being due to a built-in 'drive' released by distinctive stimuli such as the presence of another male of the same species, and inhibited by certain other signs such as a characteristic posture of submission. Lorenz transferred this theory straight over to human beings, modifying it to take account of the distinctively intercommunal nature of human carnage—which he attributed to the selective pressures of an alleged evolutionary past in which the competition for survival was more between tribes rather than individuals. If there is really such an innate tendency to communal aggression (as the bloody history of ethnic, nationalist and religious conflict suggests), then no social changes can eliminate it. The best that Lorenz can recommend is harmless redirection of it into sports, plus control by rational self-knowledge and a sense of humour.

Wilson and others offer a wider-ranging analysis of innate factors in human nature. There is less talk these days of 'instincts', as used by Freud and Lorenz, and more of a large number of genetically based predispositions which interact in subtle ways, depending on the environment, in the production of behaviour. But

113

the emphasis is still very much on innate tendencies, seen as the result of a long history of natural selection, whose detailed expression may depend on culture and individual conditioning, but which will certainly express themselves somehow or other. Yet much of what the socio-biologists say about human nature is bound to be controversial, for two reasons—because it is so difficult to separate the contributions of heredity, and because of the normative involvements which surround the topic of human behaviour. We cannot make tight connections between particular genes and identifiable kinds of social behaviour, nor is it to be expected that the science of human genetics will bring us to that stage, for there must surely remain some part for culture to play.

For example, the whole area of human sexual roles is hotly debated. Socio-biologists may point to the selective pressure on males (of all species) to spread their genes around as widely as possible, in contrast to that on females to select their partners cautiously for genetic fitness, but they also have to acknowledge that pair-bonding is a (fairly) typical feature of human behaviour, unlike the other primates. So they may try to explain both our monogamy and our frequent departures from it in terms of an evolutionary history which grafted pair-bonding (supposedly required by the hunting way of life, with its male absences from home) onto a pre-existing primate pattern of dominant male plus harem. They may try to explain traditional human sexual division of labour by our ancestors' system of males going hunting in groups (involving distinctive male-to-male bonding) while women looked after the young. But feminists, such as Alison Jaggar, resist any attempt to justify the continuation of traditional sexual roles on supposedly biological grounds; they argue that whatever may have been the case in the distant past, it is now very much a matter of culture, and is therefore challengeable and changeable.

If human nature is, at least in part, a matter of genes, then is it open to us to improve it by genetic engineering, intervening to control the very genes of future generations? This could in theory be done (somewhat slowly) by selective breeding, as the eugenics movement advocated earlier in this century (see Rose 1984)—after all, we have in this way been able to alter the characteristics of animals and plants. But perhaps when we gain knowledge of our genes themselves—the way in which they are encoded in the DNA structures of the whole human genome—it might also be done more quickly, if we find techniques to manipulate these genes at will. In both cases a distinction needs to be made between negative and positive programmes, the former aiming only to prevent the birth of physically or mentally handicapped infants, the latter trying to produce the 'best' sorts of human being. This positive selection is much more ambitious and much more controversial: which features are we to select for? Who is to decide about this: prospective parents, the state, or who? How could human reproduction be controlled in the massive way envisaged? How could

anybody have the *right* thus to interfere with other people having children? What we have here are not so much theories of human nature, but the possible means to mould it in one direction or another. Whether to use such means at all, and if so how, are questions of value.

There seems to be no escape from the conclusion that in so far as we can ascertain facts about what human nature is (and has been), this does not settle questions of value about what it ought to be. Disputed questions of philosophy and value are involved as soon as anyone tries to apply the scientific method to human nature. For there are those (philosophers, existentialists, Marxists, theologians) who in their different ways maintain that we transcend our biology–by our rationality, our consciousness, our freewill, our social development, or even our relationship to a divine Reality.

BIBLIOGRAPHY

Aristotle (*c* 300 BC) *The Politics* London: Penguin, 1951.

Freud, S. (1915–17) *Introductory Lectures on Psycho-analysis,* London: Penguin, 1973.

Hobbes, T. (1651) *Leviathan*. London: Penguin, 1968).

Jaggar, A. (1983) *Feminist Politics and Human Nature,* Brighton: Harvester.

Jones, G. (1980) *Social Darwinism and English Thought: The Interaction between Biological and Social Theory,* Brighton: Harvester.

Locke, J. (1690) *Two Treatises on Government,* London: Dent, 1924.

Lorenz, K. (1966) *On Aggression,* London: Methuen.

Marx, K. (*c*. 1843–75) *Selected Readings in Sociology and Social Philosophy,* eds T.Bottomore and M.Rubel, London: Penguin, 1963.

Plato (*c*. 380 BC) *The Republic,* London: Penguin, 1955.

Rose, S., Lewontin, R.C. and Kamin, L.J. (1984) *Not in Our Genes: Biology, Ideology and Human Nature,* London: Penguin.

Rousseau, J.-J. (1755) *Discourse on Inequality,* London: Penguin, 1984.

——(1762) *The Social Contract,* London: Penguin, 1968.

Skinner, B.F. (1953) *Science and Human Behaviour,* New York: Macmillan.

Wilson E.O. (1975) *On Human Nature,* Cambridge, Mass.: Harvard University Press.

7

CONCEPTIONS OF LEGITIMACY

MATTEI DOGAN

Why do people voluntarily follow and obey their rulers? Why do people accept and maintain authorities and institutions? In authoritarian regimes people obey involuntarily, by fear. But, as Xenophon already knew, the power of tyrants is not based uniquely on material force and constraints. Even the most tyrannic rulers try to justify their reign. The key concept to the understanding of this effort of justification is legitimacy, because only legitimacy can transform brutal power into recognized authority.

Legitimacy has always been in the mind of political thinkers. Plato's idea of justice bears on the problem of legitimacy, as well as Aristotle's distinction between monarchy, aristocracy and democracy. In his analysis of the nature of government, Locke displaced the source of legitimacy, replacing the divine right of kings by the consent of the people. No discussion of the concept of power could be complete without reference to legitimacy. For contemporary political systems in which participation of the people is a criterion of political worth, legitimacy is a fundamental concept.

DEFINITIONS OF LEGITIMACY

The concept of legitimacy and its definition have changed significantly since the emergence of democratic governments. As Schaar points out, current definitions of legitimacy dissolve legitimacy into belief or opinion (Schaar 1981). If people hold the belief that existing institutions are appropriate or morally proper, then those institutions are legitimate. Such a reference to beliefs becomes even clearer when we consider the widely accepted definition formulated by Lipset: 'the capacity of the system to engender and maintain the belief that the existing political institutions are the most appropriate ones for the society' (Lipset 1959:77). It is also clear in Merkl's definition: 'a nation united by a consensus on political values…a solemnly and widely accepted legal and constitutional order of democratic character…and an

elective government responsive to the expressed needs of the people' (Merkl 1988:21).

Juan Linz proposes as a 'minimalist' definition 'the belief that in spite of shortcomings and failures, the political institutions are better than any other that might be established, and therefore can demand obedience' (Linz 1988:65). The concept of 'diffuse regime support' developed by David Easton is another way to define legitimacy (Easton 1965).

The best-known definition of legitimacy today was formulated by Max Weber, who distinguished three types of legitimacy: traditional, charismatic and legal-rational (Weber 1978). This typology has been meaningfully applied in many historical studies: 'Since Weber, we have been busy putting the phenomenon into one or another of his three boxes and charting the progress by which charismatic authority becomes routinized into traditional authority, which…gives way in turn to rational legal authority' (Schaar 1981:15). Legitimacy is particularly important in democracies since a democracy's survival is ultimately dependent on the support of at least a majority of its citizens; it holds that at least a majority must deem it legitimate. Hence, without the granting of legitimacy by the people, a democracy would lose its authority. On the other hand, legitimacy in this sense of public belief and support is considerably less important in non-democratic regimes. In dictatorships, while the granting of support or legitimacy by the people may be an asset, it is not of ultimate importance since authority is based on force.

Authoritarian regimes may lack legitimacy but they still feel a need to acquire it. The subtitle of Michael Hudson's book on Arab politics is very significant: *The Search for Legitimacy* (Hudson 1977). He clearly explains this need:

> The central problem of government in the Arab World today is political legitimacy. The shortage of this indispensable political resource largely accounts for the volatile nature of Arab politics and the autocratic, unstable character of all the present Arab governments…Whether in power or in the opposition, Arab politicians must operate in a political environment in which the legitimacy of rulers, regimes and the institutions is sporadic and, at best, scarce. Under these conditions seemingly irrational behavior, such as assassinations, coups d'etat and official repression, may in fact derive from…the low legitimacy accorded to political processes and institutions.
>
> (Hudson 1977:2)

THE OBSOLESCENCE OF CLASSICAL TYPOLOGIES OF LEGITIMACY

In the Weberian typology, the concepts of legitimacy and democracy are not related. Historically, traditional legitimacy and charismatic legitimacy are only found in authoritarian regimes. They never appear in truly democratic regimes.

The implication is that some authoritarian regimes can be legitimate. Among the contemporary countries with legal-rational authority some are legitimate, particularly the pluralist democracies, but most are not, particularly the authoritarian regimes.

Today it is more difficult than in the past to make clear-cut classifications of authority, because the legitimacy of a regime can be based on more than one type of authority. The democracy of the United States is not based exclusively on its short, sacred Constitution. It has developed progressively, generating new practices which were soon formalized and routinized. How much rationality and how much tradition is there in the contemporary Indian democracy?

Even Max Weber has implicitly accepted this idea of mixed legitimacy. He discussed the dynamics of the process of legitimation and delegitimation (Weber 1978). The ideal types that he constructed are antagonistic only in theory. In reality, all traditional systems have some features of legality: the Chinese emperors or the Russians tsars both respected some rules of the game.

The Weberian typology is no longer helpful in the study of contemporary political regimes, because only a few countries maintain a traditional authority (for example Morocco or Saudi Arabia), and the charismatic phenomenon, so frequent between 1917 and 1980, is extremely rare today–Khomeini being the most recent example. Charismatic leadership has been replaced by a personalization of power, nourished in many cases by a cult of personality. It would be a serious mistake to confuse such an engineered idolatry with genuine charismatic leadership.

Among the 160 independent nations of the world in 1990, we can distinguish about forty pluralist democracies endowed with a legal-rational legitimacy. Even monarchies such as Britain, Spain, Belgium, Sweden, Norway, The Netherlands or Japan have a legal-rational rulership–the Crown being only a symbol. These forty countries enjoy a democratic legitimacy.

This simple account shows that two of the three Weberian types of legitimacy are almost empty, and the third one includes only one-quarter of nations. Three-quarters of all countries have authoritarian regimes deprived of true legitimacy, and consequently are not covered by the Weberian typology. In order to adapt this typology to the contemporary world, it would be necessary to add a fourth 'box' for the quasi-legitimacy type, and a fifth one for the totally illegitimate regime. There is, obviously, wide diversity among authoritarian regimes. The question here is, using Easton's terminology, how much diffuse support they enjoy.

OPERATIONALIZING THE CONCEPT OF LEGITIMACY

Scholars and politicians have the tendency to adopt the dichotomy: legitimate versus illegitimate. Since the reality is much more varied, legitimacy must come in degrees. Ranking regimes on an imaginary axis from a minimum to a maximum degree of legitimacy is a promising way for the comparative analysis of political systems. Many scholars have felt the need of such scaling: 'Legitimacy runs the scale from complete acclaim to complete rejection…ranging all the way from support, consent, compliance through decline, erosion and loss. In case of conscious rejection we may speak of illegitimacy' (Hertz 1978:320).

As Juan Linz stresses, 'no political regime is legitimate for 100 per cent of the population, nor in all its commands, nor forever, and probably very few are totally illegitimate based only on coercion' (Linz 1988:66). Legitimacy never reaches unanimity, nor do groups and individuals ever recognize equally the authority of the political power. There are apathetic popular strata and rebellious subcultures, pacifist dissidents and armed terrorists, and between these extremes many who are only partially convinced by the pretensions of legitimacy claimed by the rulers. The support of the majority is generally considered as a test of legitimacy, but as David Easton observed, it is also necessary to consider the substance and intensity of the popular support (Easton 1965).

Easton argues that the 'ratio of deviance to conformity as measured by violation of laws, the prevalence of violence, the size of dissidence movements or the amount of money spent for security would provide indices of support' (Easton 1965:163). But it is difficult in empirical research to measure 'violations of laws' or 'dissident movements'.

Thus we should not assume that in a given country legitimacy exists simply because it is not contested. In the poorest countries the problem of illegitimacy is not present in the mind of the majority of the people. In these countries tyrants are often perceived as a fatality. Where violence is absent, legitimacy is not necessarily present. The concept of legitimacy is not adequate for, perhaps, one out of every five Third World countries.

Absence of revolt, however, does not imply adhesion to the regime. Revolt is possible only in certain historical circumstances, when a regime starts a process of liberalization. In a totalitarian regime attempts to revolt can be suicidal. The Chinese communist establishment, by repressing the demonstrations in the Tienanmen Square in June 1989, wanted to stop the incipient liberalization movement.

The number of *coups d'état* is the most visible measure of illegitimacy: look for instance at *coups* in Africa in the last three decades, and earlier in Latin America. This criterion has been adopted by a number of scholars.

119