

*Dieter Birnbacher*

## **Binmore's Humeanism**

*Abstract:* David Hume is quoted in Binmore's book *Natural Justice* more than any other author, past or present, and throughout with a markedly positive attitude. It is argued that this affinity is reflected in many characteristic features of Binmore's approach to fairness and social justice and especially in the central role motivational issues are made to play in his theory. It is further argued that Binmore shares with Hume not only important strengths but also certain weaknesses, among them a tendency to derive from the limited evidence of past history far-reaching statements on human nature and the conditions thereby imposed on social morality.

For a work of a theoretical economist and game theorist, Binmore's *Natural Justice* is a strikingly unorthodox book. It brings some a priori work to bear on the topics with which it deals, but on the whole its spirit is much more Humean than Kantian. This is true not only because David Hume is quoted in this book more than any other author, past or present, and with a markedly positive attitude. It is true also because it discusses the evolution of social rules and the dynamics of social justice and morality by the same combination of perspectives that Hume took toward all topics he dealt with: empiricism, skepticism, and a strong interest in the psychological and anthropological origins of the phenomena studied. In Hume's empiricist spirit, morality is made the object of a strictly empirical inquiry. Social morality is approached from 'outside', with the eyes of the historical sociologist looking at the great variety of moralities of different cultures and times and searching for a 'deep structure' that seems sufficiently universal to risk the hypothesis that the core of morality is hard-wired and encapsulated in our genes as a common biological heritage. This hypothesis seems plausible especially after Franc de Waal and others have opened our eyes about the extent to which altruistic human behavior patterns that we would intuitively call 'typically human' are present also in chimpanzees. It appears that the roots of morality are much older than the human species, although evolution had to wait for the advent of *Homo sapiens* to give morality the cultural variety we find in written and unwritten history.

There is also a good deal of Humean skepticism in the book, not only concerning the extravagant claims of some philosophers about the existence of eternal values, but also concerning its own hypotheses that are throughout put forward as no more than pieces of intelligent speculation. And there is in Binmore's book a pervasive interest in what may be called the 'motivation problem', the problem how fairness and other kinds of moral behavior can be explained in the framework

of a realistic anthropology that recognizes the primacy of egoism in the common run of human affairs. As one should expect from a professional economist, one of Binmore's central criticism of most standard theories of normative ethics is a profound psychological skepticism about the extent to which moralities can be expected to function that essentially rely on genuinely altruistic motivations instead of self-interest mitigated by what C. D. Broad and John Mackie called 'self-referential altruism', the emotional ties that bind us to family and friends. The evidence of everyday life and of history strongly suggest that human altruism is severely limited, which means that, in general, political reforms favoring one group will be acceptable to other groups only if some incentive is provided for them, at least in the absence of an external authority.

It is one of the central, though largely implicit messages of Binmore's book that the term 'incentive' must not be interpreted too narrowly. The incentive that makes one party accept a reform that favors another party need not be material or a gain in reputation or self-respect. It is sufficient that those that have to bear the costs of reform perceive the reform as *fair*. But what does fairness mean in this context? Fairness is, for Binmore, an umbrella term that cannot be tied down to any particular conception of justice but which covers an open-ended variety of culture specific norms of what is seen as right and proper or, as Binmore prefers to say, 'seemly'. Fairness in this sense is not, as it is in Rawls, an inherently liberal concept. If I understand Binmore's use of the term correctly, the content of this concept is so broad that it does not even exclude honoring privileges based on traditional conceptions of social worth or traditional understandings of role and status (which constitute so many variants of 'social indices' by which the gains or losses of a group are weighted). Fairness, as used by Binmore, is therefore more a principle of conservatism than a principle of reform. He rightly reminds morally motivated social reformers to pay more attention than they usually do to the fairness norms of those groups on which the costs of the reform fall. But since the content of these fairness norms is deliberately left open, nothing excludes that these norms appear indefensibly unjust from a liberal point of view. This, again, is fully in line with Hume's moral descriptivism. Like Hume, Binmore is more interested in questions of feasibility than in questions of justification. Indeed, he at one point goes so far to suggest that once questions of feasibility have been satisfactorily answered there is not very much else to do for the consultant giving advice in political matters: "To make a decision sensibly, first decide what alternatives will work, and then choose whichever of the workable alternatives you like most." (199)

But there are more affinities between Binmore and Hume than these methodological ones. Both are authors of profoundly good sense, relying on their own perception and with a healthy distrust of false authority. Both are authors who immediately invite not only empathy but also sympathy. And both are authors of exquisite stylistic abilities, writing on abstract matters in a remarkably lively way. Where there is light, however, there is also shadow. Both authors share not only strengths but also weaknesses. One of these is their indulging in polemics against metaphysics and especially against metaphysical ethics. This polemics, it seems to me, is much more excusable in Hume than in Binmore.

Hume's polemics against metaphysical ethics was more justified in a period in which ethical Platonism was the ruling creed and in which disrespect of theological doctrine was regarded as perfectly adequate for denying a chair even to the greatest philosopher of his times. Binmore's polemics seems much less to the point. Nowadays, moral realism is defended only by a minority of ethicists, and most versions of it are watered down to an extent that it is hard to see why they provoke the harsh criticism Binmore levels at them. On the whole, appeals to a supernatural origin of morality are much less common among moral philosophers than they are among politicians and leader writers in conservative newspapers who exploit theological fictions to give accepted social values a halo of supernatural validity. If there is anyone to blame for inventing "sources of absolute authority" (19) it is politicians who introduce a reference to God into constitutions for the simple reason of seemingly investing the values incorporated in them with some mysterious kind of 'higher' authority (as the Land Hessen in Germany did a few years ago). What should be criticized here is not the bad philosophy behind this political move but its intellectual dishonesty. Differently from the political class of the United States, only few members of the German political class are Christian fundamentalists who believe that moral values come directly from God instead of being the products of a long and circuitous process of biological and cultural evolution. What should be blamed is the unabashed instrumentalism of this political move. It is a betrayal of the Enlightenment tradition not so much because of the unenlightened beliefs that are instrumentalised than by the arrogation of a right to Platonic white lies.

The other feature that connects Binmore's approach to social affairs to Hume's, and with which I am not entirely happy, is the tendency to draw far-reaching ethical conclusions from a conception of human nature derived from the admittedly very imperfect performance of mankind so far. Though Binmore's anthropological tenets are, as I said, presented in an empiricist spirit and put forward as hypotheses rather than essentialist dogmas, they nevertheless play a crucial role in Binmore's discussion of the original position and of the conclusions drawn from it by Harsanyi and Rawls. I think that Binmore is right in stating that Rawls draws the wrong conclusions from his model by making use of a decision criterion that is evidently inappropriate in the original position. Nevertheless, he welcomes the egalitarian conclusions Rawls derives from this misapplication, though primarily for pragmatic rather than ethical reasons. Though, if the model of the original position is followed, a utilitarian conception of fairness comes out as the 'first-best creed', this creed is, given the imperfections of human motivation, seriously deficient if used as a practical strategy. Even if no more than a 'second-best creed' from the perspective of the original position, an egalitarian conception of fairness is much better adapted to real life conditions. Given that there are only two kinds of social contract to compare, anthropological evidence stemming both from anthropological data about hunter-and-gatherer societies as well as from human history and everyday experience suggest that a utilitarian strategy would inevitably run risks an egalitarian strategy might effectively avoid.

First, it seems to be part of human nature to have a clear preference for egalitarian distributions even if they are less than optimal on utilitarian terms. Second, egalitarian strategies are in general more acceptable in a democracy in which political decisions ultimately rely on the acceptance of these decisions by the whole, or at least a substantial part of the population. Whereas a utilitarian strategy might imply that some members of society are forced to make heavy sacrifices for the over-all well-being of society as a whole, an egalitarian strategy often implies that sacrifices are unnecessary because of the higher value attributed to equality over efficiency. This is one reason why an egalitarian strategy must be expected to be much more stable than a utilitarian one. Stability, however, is a central condition of feasibility. Social contracts must be stable if they are to survive, and they must be efficient under non-ideal conditions. Third, one of the central weaknesses of utilitarianism is that it is inevitably bound up with the idea of an enlightened philosopher-king who enforces a utilitarian morality. It requires a central agency able to "provide all players with the right incentive to do just the right thing at the right time" (92). This is, as the history of mankind shows, a utopian ideal that cannot be expected to work in practice. Instead of being dependent on an external agency, a social contract that can be expected to function under real-life conditions must be self-supportive.

Each of these anthropological arguments is well taken, though I doubt that they are sufficient to make a convinced utilitarian ethicist change his mind. As pragmatic arguments they are worth being borne in mind by anyone who intends to put his ethical principles into practice. But as such they are certainly not sufficient to determine the ethical stance one takes regarding questions of social justice. It must be conceded that the anthropological evidence adduced by Binmore for the hypothesis that a preference for egalitarianism is written into our genes is impressive. This hypothesis might explain the nearly automatic association between the ideas of distributive justice and the idea of equality evidenced by psychological studies.

It might explain, furthermore, the fact that, on the whole, egalitarian societies are both those in which people feel happiest and which exhibit a very high level of productivity. All this, however, implies only that a follower of Harsanyi on the ethical level should possibly nevertheless act as a Rawlsian on the level of practical policy. If equality is one of the things that makes people *really* happy (as I think it does), a utilitarian ethical strategy should be thoroughly egalitarian on the level of politics. And the same holds for the two other arguments. On the level of strategy, stability is of primary importance. If so, the long route from the status quo to a distant social Utopia, however defined, must be regulated by principles that can be expected to be acceptable to sufficiently many people to attain its desired end. Constructions like the original state tend to obscure this point because they focus only on result quality and not also on process quality. But on utilitarian as on egalitarian principles, not only the state in which the respective ideal is realized but also the ways to achieve this must stand the ethical test. Not only the ends, also the means must be ethically defensible.

And this, in general, requires the means to be consensual rather than externally enforced. Far from being somehow analytically tied to some form of

moral dictatorship, utilitarianism strongly favors internal control mechanisms. Any morality needs an authority that enforces it, but this need not be thought of as an agency external to the society or group in which this morality is enforced. The authority can just as well be internal to the group, for example in the shape of an established mechanism of mutual social control, as in small religious communities. Binmore mentions the possibility of ‘strong peer pressure’ as an alternative to external authority in his opening paper (though, as far as I see, not in the book). Since the utilitarian knows perfectly well that people “are happier when they aren’t bossed around all the time” and that they tend to be “immensely more productive” (189) without the frustrations of obedience, the utilitarian strategy of reform will only in rare cases result in a strictly hierarchical welfare state on the model of an enlightened despotism. Moreover, there are, as John Stuart Mill saw, good utilitarian reasons both for civic and personal freedoms and for incentives for self-initiative and individuality.

This amounts to saying that anthropological arguments are, in a way, necessary but not sufficient for deciding on ethically sensitive issues like fairness and social justice. Moreover, it is always reasonable to ask whether the conditions under which certain principles of human behavior have evolved are sufficiently similar to the conditions under which the same principles are applied in a concrete case to suggest the same conclusions. In a famous passage of his *Treatise*, Hume wrote that

“in general, it may be affirm’d, that there is no such passion in human minds, as love of mankind, merely as such, independent of personal qualities, of services, or of relation to ourself”.

This is true as far as it goes. However, in the next sentence Hume makes a concession that might prove to be of some importance in an age in which the question with whom and to what extent we empathize is largely decided by the way events are represented in the media: “’Tis true, there is no human, and indeed no sensible, creature, whose happiness or misery does not, in some measure, affect us, when brought near to us, and represented in lively colors.” Though human nature may not be made for global sympathies, an ethics of international solidarity, which was, in Hume’s day, an outright anthropological impossibility, might well prove to become a real possibility under changed external conditions.

## Bibliography

Binmore, K. (2005), *Natural Justice*, Oxford-New York