

Review of

Ken Binmore, *Natural Justice* (Oxford: Oxford University Press, 2005), pp. xii + 207

Review by Karl Widerquist

This is an early version of a review that was eventually published as 2009. "Review of *Natural Justice*, Ken Binmore," *Utilitas* 21 (4): pp. 529-532. Please refer to the published version

Ken Binmore's new book is simultaneously an introduction to game theory, a lesson in social engineering, and a treatise social justice inspired by Hume and the Whigs. These three objectives fit into one book because Binmore equates justice with a working social contract and uses game theory to examine what sorts of social contracts work. As an introduction to game theory, and as a lesson in social engineering, the book is very good. As a treatise on social justice, it is less than convincing.

Binmore has a great ability to explain game theory for non-specialists, but he uses this ability inconsistently. He creates very accessible examples to show why game players end up at a Nash equilibrium, but he does have a similarly accessible example to illustrate the Nash bargaining solution. The book is still too technical to achieve his goal of writing a popular science version of ideas laid out in his two-volume work, *Game Theory and the Social Contract*. He mentions in the preface that the book contains no algebraic equations, but it does contain several *arithmetic* equations and some graphs and tables that are as technical as those in his earlier work.

By *Natural Justice*, Binmore means that justice and fairness can only be understood as part of the process of human physical and cultural evolution. Every society needs a set of common understandings to help individuals coordinate their efforts. He calls this set a social contract, and examines how to engineer a good one. That is, one with three features—stability, efficiency, and fairness. He writes that most political philosophers completely ignore the first two features, and usually get the third one wrong.

The most important thing Binmore has to say, can be summarized in three sentences: No social contract will succeed unless it is both stable and efficient. Any project to organize society that ignores these features it is therefore folly. But there are many different social contracts that are both stable and efficient, and from that set, we can select whatever social contract we find to be fair, just, or desirable. Whether or not one agrees with Binmore's social philosophy, this book is worth reading just for his discussion of how to ensure social contracts are stable and efficient.

According to Binmore, too many philosophers and social theorists have assumed that once their preferred social contract is in place, people will miraculously behave as designers wish. Game theory can help us identify stable social contracts by identifying the equilibrium reached when individuals play their optimal strategy under a given social contract. To use it, we need a realistic assessment of how people actually behave. Binmore argues that people behave mostly selfishly, punctuated by reciprocal altruism in accordance with prevailing fairness norms. After looking back at the book several times, I am still unclear to what extent fairness affects how people behave or what social contracts they will accept. He believes that people will act to enforce reciprocal altruism—help those who help others. But he does not believe we can count on people to sacrifice their own well-being for the common good. Instead, we have to set the rewards so that it is in everyone's self-interest to do things that promote the common good.

Binmore has good advice for any social planner. Don't count on people to behave against their self-interest, but design a mechanism so that we reach a stable, efficient, and fair outcome even though everyone follows their own self-interest.

But this is not the only limit on social planners. They also have to follow prevailing norms about what is fair or just, not only because people will not accept a social contract with goals too far from those that most people believe to be right, but also because Binmore endorses a strong form of cultural relativism. To Binmore justice is simply the fulfillment the prevailing social contract. While most philosophies strive for the good or the right, Whigs strive for the seemly. Whatever society can usually effectively punish is immoral; everything else is permitted. To say that one social contract is better, worse, or equal to another is meaningless. This belief limits us from condemning the practices of another society working on a different social contract, but it frees us to move from one social contract to another at will.

Binmore's defense of cultural relativism is the weakest part of the book. As he sees it, there are only two possibilities. Either a person recognizes that her beliefs about justice are affected by culture and endorses cultural relativism, or a person denies cultural relativism and asserts her own beliefs are unaffected by culture. There are other possibilities that this book ignores. A person could recognize that her own beliefs are affected by culture while simultaneously believing (1) there is a universal morality but no one know for sure what it is; (2) there is no such thing as morality; or (3) there is no way to tell whether morality is relative, universal, or nonexistent.

Binmore uses the example of a change in traffic laws. Before 1967, it was immoral to drive on the right in Sweden; then it became moral. To say that society should change the rule is not to say that driving on the right is more moral than driving on the left; it is merely to say that one wants it to become moral to drive on the right. According to Binmore's relativism, driving on one side is not moral or immoral because it leads to more people getting killed. That would be a universal moral standard. It is moral or immoral because society can effectively punish a driver who deviates from that rule.

Binmore lacks the courage of his convictions on cultural relativism. He writes that, if he lived in Aristotle's time, he would have *thought* slave holding was moral. If he lived in Spain during the inquisition or in Germany during World War II, he would have *thought* it was right to brutalize Jews. If lived in Victorian times, he would have *thought* it was moral to disown a daughter who gave birth to an illegitimate child. If morality is the fulfillment of a social contract, it is not enough to say he would have *thought* these things were moral; he has to say that these things *actually were moral*.

He defends cultural relativism against the charge that it implies moral subjectivism (that anything goes) on the grounds that the function of moral rules is to help people coordinate their behavior, which would be impossible if everyone followed their own rules. This defense does not cover situations in which no effective social contract exists. During a riot, in international relations, or in places where conflicting social overlap, there is no effective contract to break, and anything *does* go, whether it is robbery, murder, or genocide. This argument does work to keep morality from becoming relative at the individual level, but by doing so, it has the undesirable implication that an individual is immoral to subvert a social contract that she never endorsed and that does not benefit her. A slave who tries to escape is objectively *immoral*.

Binmore also defends cultural relativism against the charge that it implies resisting any reform, by arguing that relativists can advocate for whatever reform they like simply because they like it. This statement is true only if the relativist lives in a society that protects freedom of speech, but not in a society like China today or the American south during slavery. We are mistaken to think that people who advocate for reforms at great risks to themselves are brave and moral. It is an objective fact that they are immoral in the same way and for the same reason that pedophiles and people who drive on the wrong side of the road are immoral in our society today.

Binmore is vague about what reforms he would actually like to see. He wants some kind of incremental, sustainable reforms in a Whiggish direction. But I would argue that there is no such a thing in as a Whiggish direction. If Whigs can advocate any stable, efficient equilibrium they like, all small changes in any direction are equally Whiggish. He argues that humans may have a predisposition to return eventually to the egalitarian social contracts of our hunter-gatherer ancestors, but he is aware that hierarchical social contracts have been often managed to find stability. He states that a Whiggish utopia might look something like the early years of the American Republic. By this, I assume he does not mean slave holding and wife beating will again become moral, but whatever he does me is left unsaid.