Analyse & Kritik 01/2012 (© Lucius & Lucius, Stuttgart) S. 129–139

Ken Binmore

Kitcher on Natural Morality

Abstract: This commentary on Philip Kitcher's *Ethical Project* compares his theory of the evolution of morality with my less ambitious theory of the evolution of fairness norms that seeks to flesh out John Mackie's insight that one should use game theory as a framework within which to assess anthropological data. It lays particular stress on the importance of the folk theorem of repeated game theory, which provides a template for the set of stable social contracts that were available to ancestral hunter-gatherer communities. It continues by drawing attention to the relevance of Harsanyi's theory of empathetic preferences in structuring the fairness criteria that evolved as one response to the equilibrium selection problem that the folk theorem demonstrates is endemic in our species.

1. Introduction

Ever since David Hume inadvertently woke Immanuel Kant from his dogmatic slumber, moral naturalism has been a poor relation at the philosophical table, seated well below the salt and regarded as a vulgar interloper unacquainted with the finer things of life. Naturalists without formal philosophical qualifications, like Robert Sugden (1986) or myself, might as well not exist at all. But now that Hume is being taken seriously again, naturalists may reasonably live in hope that a tipping point has been reached, and that all that is necessary for naturalism to get the hearing it deserves is for a respected philosopher to come along who knows how to persuade traditional moral philosophers to turn on their hearing aids.

Is Philip Kitcher the fairy godmother we have been waiting for? Will he succeed where such predecessors as John Mackie's (1977) *Inventing Right and Wrong*, Peter Singer's (1980) *Expanding Circle*, Brian Skyrms' (1996) *Evolution of the Social Contract*, and Robert Nozick's (2001) *Invariances* seem to have got nowhere at all? I don't know the answer. But then I don't understand why Mackie's devastating critique of the traditionalist position should be treated with an indifference that amounts almost to contempt. So my opinion on whether Kitcher's notion of *pragmatic naturalism* (in honor of John Dewey) is likely to hit the spot with mainstream philosophers is not worth very much. I hope I shall therefore be forgiven for restricting my comments to a comparison of Kitcher's speculations on the evolutionary origins of human morality with my

own less ambitious, but more specific, writings on the evolution of fairness norms (Binmore 1994; 1996; 2005). To organize the issues, it is useful to recall the standard personification of justice as a blindfolded matron bearing a pair of scales and a sword.

2. The Sword

The sword is commonly neglected in traditional discussions, but as Hobbes famously observed: covenants without the sword are but words.

Anthropologists seem to agree that the origins of human sociality are to be found in the food-sharing arrangements to be found in many species, whereby animals succeed in insuring each other against hunger. But what induces Adam to honor an implicit deal to share food with Eve on those occasions when she is hungry and he has food to spare? David Hume offered an answer in 1739:

> "I learn to do service to another, without bearing him any real kindness, because I foresee, that he will return my service in expectation of another of the same kind, and in order to maintain the same correspondence of good offices with me and others. And accordingly, after I have serv'd him and he is in possession of the advantage arising from my action, he is induc'd to perform his part, as foreseeing the consequences of his refusal."

When the biologist Robert Trivers (1971) reinvented this notion, he had the brilliant idea of calling it *reciprocal altruism*. Kitcher is fulsome in his praise of Trivers (and also of Axelrod and Hamilton), but actually none of these scholars was the first to reinvent Hume's insight. Its modern incarnation dates from the early 1950s, when John Nash was asked by Dresher and Flood to comment on the results of their very first experiment on the Prisoners' Dilemma, in which the same two subjects played each other 1,000 times. Nash pointed out that repeated games have more Nash equilibria than that obtained by simply repeating a Nash equilbrium of the game that is repeated. A number of researchers (notably Robert Aumann) then took this observation further by formulating a theorem that describes *all* possible equilibrium outcomes of an infinitely repeated game played between infinitely patient players whose history of past play is always available for inspection. Since nobody knew to whom to attribute the result, the theorem became known as the *folk theorem*. I think an understanding of this result is a *sine qua non* for political philosophy, but some preliminary discussion of elementary game theory is necessary to explain it.

Games and equilibria. My *Very Short Introduction to Game Theory* (Binmore 2007) is an equationless introduction to the basic ideas of game theory, but philosophers may prefer one or other of Brian Skyrms' very readable books (Skyrms 1996; 2003).

A game is a way of formalizing any situation in which people or animals interact with each other. The basic idea in analyzing a game is that of a Nash equilibrium (Nash 1951), which is a strategy profile—one for each player—in

which each player's strategy is a best reply to the strategies chosen by the other players. Nash equilibria are of interest for two reasons. If it is possible to single out the rational solution of a game, it must be a Nash equilibrium. For example, if Adam knows that Eve is rational, he would be stupid not to make the best reply to what he knows is her rational choice. The second reason is even more important. An evolutionary process that adjusts the players' strategy choices in the direction of increasing payoffs can only stop when it reaches a Nash equilibrium.

Because evolution stops working at an equilibrium, biologists say that Nash equilibria are evolutionarily stable.[1] Each relevant locus on a chromosome is then occupied by the gene with maximal fitness. Since a gene is just a molecule, it cannot *choose* to maximize its fitness, but evolution makes it seem as though it had. This is a valuable insight, because it allows biologists to use the rational interpretation of an equilibrium to predict the outcome of an evolutionary process, without following each complicated twist and turn that the process might take.

The title of Richard Dawkins' (1976) *Selfish Gene* expresses the idea in a nutshell, but it also provokes a lot of criticism. It is easy to be tolerant of critics like the old lady I heard rebuking Dawkins for failing to see that a molecule cannot possibly have free will, but tolerance is less easy in the case of critics like Lewontin or Gould, who chose to whip up public hostility against Edward Wilson and his followers on similar grounds. As Alcock's (2001) *Triumph of Sociobiology* documents, they willfully pretended not to understand that sociobiologists seek explanations of biological phenomena in terms of *ultimate* causes rather than *proximate* causes. Why, for example, do songbirds sing in the early spring? The proximate cause is long and difficult. This molecule knocked against that molecule. This chemical reaction is catalyzed by that enzyme. But the ultimate cause is that the birds are signaling territorial claims to each other in order to avoid unnecessary conflict. They neither know nor care that this behavior is rational. They just do what they do. But the net effect of an immensely complicated evolutionary process is that songbirds behave *as though* they had rationally chosen to maximize their fitness by operating a Nash equilibrium of their game of life.

Prisoners' Dilemma. The Prisoners' Dilemma is the most famous of all toy games. A whole generation of scholars swallowed the line that this trivial game embodies the essence of the problem of human cooperation. The reason is that its only Nash equilibrium calls for both Adam and Eve to refuse to cooperate, but they would both get more if they did cooperate. The hopeless task that scholars set themselves was therefore to give reasons why game theory's resolution of this supposed 'paradox of rationality' is mistaken.

Game theorists think it just plain wrong to claim that the Prisoners' Dilemma embodies the essence of the problem of human cooperation. On the contrary, it represents a situation in which the dice are as loaded against the emergence

[1] John Maynard Smith (1982) defines an evolutionarily stable strategy (ESS) to be a best reply to itself that is a better reply to any alternative best reply than the alternative best reply is to itself, but biologists seldom worry about the small print involving alternative best replies.

of cooperation as they could possibly be. If the game of life played by the human species were the Prisoners' Dilemma, we would not have evolved as social animals! We therefore see no more need to solve some invented paradox of rationality than to explain why strong swimmers drown when thrown in a lake with their feet encased in concrete. No paradox of rationality exists. Rational players do not cooperate in the Prisoners' Dilemma, because the conditions necessary for rational cooperation are absent in this game.

Fortunately the paradox-of-rationality phase in the history of game theory is long over. Insofar as they are remembered, the many fallacies that were invented in hopeless attempts to show that it is rational to cooperate in the Prisoners' Dilemma are now mostly quoted as entertaining examples of what psychologists call magical reasoning, in which logic is twisted to secure some desired outcome. The leading example remains Kant's claim that rationality demands obeying his categorical imperative. In the Prisoners' Dilemma, rational players would then all choose to cooperate because this is the strategy that would be best if everybody chose it.

If one wants rational players to cooperate in the one-shot Prisoners' Dilemma (as opposed to a repeated version of the Prisoners' Dilemma), it is necessary to provide them with some means of external enforcement, such as a judicial system. The players can then sign a *binding* contract in which they agree to cooperate. Neither player will then wish to cheat on their contractual obligation, because they will then be punished by the external enforcement agency. Without some means of external enforcement, any agreements between the players need to be self-policing—and the only *self-policing* agreements are agreements to operate one of the Nash equilibria of whatever game is being played. Nobody has an incentive to be the first to deviate from the terms of a contract that requires everybody to operate a Nash equilibrium, because all players are already doing as well as they can given the behavior of the other players in the game.

In particular, the implicit insurance contract built into a food-sharing arrangement needs to be a Nash equilibrium of whatever game of life the animals who operate it are playing. But animals operating a food-sharing arrangement are never playing a *one-shot* game like the Prisoners' Dilemma. They are always playing a repeated game in which the kind of reciprocity Hume was talking about has a chance to get off the ground.

Folk theorem. The folk theorem says that all outcomes of a one-shot game on which the players might like to sign a binding agreement if an external enforcement agency were present are available as Nash equilibrium outcomes in indefinitely repeated versions of the game—provided that the players have no secrets from each other and care enough about the future for long-term relationships to outweigh short-term benefits. The no-secrecy condition does not fit modern societies very well, but it must be remembered that we are discussing the emergence of moral norms in the small hunter-gatherer communities that preceded the agricultural revolution of some ten thousand years ago.

The proof follows the lines proposed by Hume. The essential idea is that any deviation from a contract will be followed by some kind of punishment by the other players. It follows that *any* outcome of the one-shot game that assigns each

player more than the worst punishment that the other players can inflict on him is available as a Nash equilibrium outcome in the repeated game. No outcomes on which the players would want to write a contract are therefore excluded.

The punishment may simply consist of withdrawing future cooperation as Hume suggests, but it could also be something much more pro-active. In his usual prescient style, Hume points out that if one player deviates by victimizing another, the punishment need not be administered by the victim, but by other players—a point commonly overlooked by those social scientists who think that all one needs to know about reciprocity is encapsulated in the strategy TIT-FOR-TAT.[2]

Hume (1739) also understood that the number of players is important, because it is much easier for people to keep deviant behavior secret in a large society:

> "Two neighbours may agree to drain a meadow, which they possess in common; because 'tis easy for them to know each other's mind, and each may perceive that the immediate consequence of failing in his part is the abandoning of the whole project. But 'tis difficult, and indeed impossible, that a thousand persons shou'd agree in any such action."

Elinor Ostrom won her Nobel Prize in 2009 partly for exploring the political institutions that societies have evolved for coping with the monitoring problem in large societies that I think Hume must have been the first to identify in such clear terms. But no such monitoring problem was present in our ancestral hunter-gatherer societies.

Kitcher on enforcement. It is frustrating for a game theorist to find no mention of Nash equilibria or the folk theorem in Kitcher's book. The biological literature he quotes is similarly barren, presumably because the early game theory literature was quite unnecessarily mathematical. But the neglected ideas seem to me essential tools for any attempt to understand how human societies work. In particular, it is not adequate to substitute the biological notion of an evolutionary stable strategy (ESS) for that of a Nash equilibrium for at least two reasons. The first is that perfectly respectable evolutionary processes can easily converge on strategies in symmetric games that are *not* ESS (Hofbauer/Sigmund 1998). The second is that *no* pure strategy at all can be an ESS in a generic repeated game.

Kitcher does give some simple examples of games in which cooperation is achievable as an equilibrium outcome, but the time has gone when such examples could be regarded as adequate, because we now know that one can make all kinds of things happen by varying the conditions of play. Axelrod's (1984) claims for the strategy TIT-FOR-TAT provide a salutary example. Kitcher repeats the much quoted claim that TIT-FOR-TAT won Axelrod's evolutionary competition, but even this is not true. It was merely slightly more numerous at the end of his

[2] The strategy says to start by cooperating in the repeated Prisoner's Dilemma, and thereafter to copy whatever the opponent did in the previous round (Axelrod 1984).

evolutionary simulation than five other surviving strategies, each of which was being played by about one sixth of the population.[3]

3. The Scales

Justice needs her pair of scales to weigh up the welfare of different people when judging what is fair. Behavioral economists believe that their experiments demonstrate that people have social or other-regarding utility functions into which such interpersonal comparisons are incorporated.[4] Most people certainly care about their near and dear, and the fact that most of us contribute some small fraction of our income to charity suggests that we care to some extent about strangers too. However, I am doubtful that we care about strangers to the extent claimed by the behavioral school (Binmore/Shaked 2009). But these doubts are irrelevant to my own moral theory, which takes the personal utility functions of the citizens of a society as given, whatever they may be. That is to say, it works whether people are assumed to be Jekylls or Hydes.

Emotional altruism? The issue is more important to Kitcher, who makes the evolution of social preferences central to his approach, although he does not seem aware of the behavioral economics literature that would support his case. To illustrate what a social utility function is, he offers a formula in which a person is assumed to seek to maximize a weighted sum of his own welfare and that of a companion. The weights determine the extent to which the person values his own welfare above or below the welfare of his companion. A linear formula of this kind echoes William Hamilton's famous notion of *inclusive fitness*, according to which animals ought to care about their relatives in proportion to their degree of relationship.[5]

If I understand him, Kitcher regards the kind of caring that results from the use of Hamilton's rule within the family as an example of *biological altruism*. However, he argues that such biological altruism is inadequate to explain human morality, with which I do not imagine anyone will disagree. But he goes much further in denying that *behavioral altruism*, in which people treat strangers *as though* they cared about their welfare is adequate either. He therefore rejects the kind of reciprocal altruism on which my theory relies. He insists that what is

---

[3] In fact, all the claims that Axelrod makes for the specific strategy TIT-FOR-TAT fail to survive serious scrutiny (Binmore 2001). In particular, TIT-FOR-TAT is obviously not an ESS in the repeated Prisoners' Dilemma, because a population consisting only of players using TIT-FOR-TAT can be invaded by players who always cooperate no matter what.

[4] They usually assert that these claims refute homo economicus as a model of man because they show that people are not selfish. Kitcher shares this common misunderstanding about the nature of homo economicus, which presumably derives from the assumption in neoclassical economics that agents maximize their own utility functions. However, nothing says that their utility functions need be selfish. For example, it is always taken for granted that heads of households care about the welfare of their families. In fact, economics students are traditionally taught: *De gustibus non est disputandum.*

[5] For example, I should care about a full cousin one eighth as much as I care about myself because the probability that we have the same gene at any locus on a chromosome is 1/8.

needed is not even psychological altruism, but what he calls *emotional altruism*, in which people really do care about their fellow men deep inside.

I have two problems with Kitcher's insistence on emotional altruism. The first is methodological. I think it unwise in an evolutionary discussion to insist on particular proximate mechanisms before the ultimate issues have been sorted out. Obviously our emotions are heavily involved in our moral intercourse, but exactly how they are involved is something that needs to be studied *after* we have fully explored the evolutionary function that they serve. The second problem is standard in evolutionary discussions. How can animals who are somehow programmed to care about the fitness of others without expectation of any compensating fitness benefit survive the appearance of cheaters who behave as though they are equally nice but actually behave selfishly when they think they can get away with it? Or to put the same point another way, how can a strategy profile that is not in equilibrium survive in the long run?

The game theory literature contains numerous papers that seek to explain how people may come to behave as though they care for others beyond the extent to which Hamilton's rule applies. Although they express themselves in different ways, the successful arguments all boil down to the necessity of there being some *correlation* between the fitnesses of the players involved, through assortative mating or some other natural mechanism. Only then can one hope for such caring behavior to be evolutionarily stable. Space does not allow a survey of this literature, so I shall just mention that the latest paper of which I know not only contains a new model of this kind but also provides a large bibliography (Alger/Weibull 2012).

However, Kitcher (*EP*, 112) seems to insist that moral codes can be so internalized that neither the carrot of correlated fitness nor the stick of real punishment is necessary for their maintenance. He goes on to offer the invention of supernatural beings as a possible substitute for the latter. However, I am with David Hume (1777) on this:

> "Hear the verbal protestations of all men: Nothing so certain as their religious tenets. Examine their lives: You will scarcely think that they repose the smallest confidence in them."

My own unpopular view is that a moral naturalist has to put aside the idea that morality somehow evolved as a *substitute* for power. It evolved instead as a device to allow human beings to coordinate their behavior on one of the many ways of *balancing* power. This claim is usually interpreted as an evil economist's invitation to a spree of unbridled excess, but the critics who take this line fail to recognize the breadth of the set of equilibria available to a society. There are always very inefficient equilibria in which the citizens fight and squabble, but the folk theorem implies that any such equilibria are necessarily dominated by efficient equilibria, in which the citizens share and cooperate without the need of postulating consciences made of iron or the fear of supernatural vengeance.

Kitcher and I offer essentially the same group selection argument[6] for why evolution should be expect to eliminate societies with inefficient social contracts

[6] Kitcher prefixes his remarks on this subject with a discussion of the importance of coali-

in favor of societies with efficient social contracts.[7] The citizens of efficient societies will have more children on average and hence will spawn daughter societies that take the social contract of their parent society to new locations more often. However, my version of the story is immune to the standard criticism of group selection stories because the fact that I insist that stable social contracts must be equilibria means that no individual can profit by being the first to deviate from what is good for the group.

But finding an evolutionary reason why social contracts should be expected to be efficient leaves a large problem to be resolved. There are always *many* efficient equilibria. How is evolution to solve this equilibrium selection problem? I think fairness norms evolved in the human species for this purpose—not after the evolution of language as Kitcher would probably argue—but before or during the evolution of language.

Empathy and sympathy. I return now to the subject of interpersonal comparison to draw attention to the work of John Harsanyi (1977). One may not like Harsanyi's use of these ideas in defending utilitarianism, but they are equally useful in discussing other theories of the social contract—notably the kind of egalitarianism proposed by Rawls (Binmore 2005).

Psychologists distinguish between empathy and sympathy. The social preferences postulated by Kitcher—in which people feel the happiness or distress of others in some degree as though it were their own—come under the heading of sympathy. Empathy is the facility to put oneself in the position of others to see things from their point of view. A confidence trickster may therefore empathize with an old lady with a view to concocting a story that will persuade her to part with her savings without feeling any sympathy for her subsequent plight at all.[8]

Following Patrick Suppes (1966) and others, Harsanyi (1977) applies standard utility theory to what I call empathetic preferences.[9] Such preferences are to be distinguished from the personal preferences (that may be sympathetic in character) which people reveal when they make economic choices. By contrast, empathetic preferences are revealed, for example, when somebody says that she would rather be Adam drinking a cup of coffee than Eve drinking a cup of tea. Harsanyi shows that if a person is sufficiently adept at empathetic identification that she agrees that, if she were a particular person, then her personal preferences would be the same as that person's, then her empathetic utility function will be a weighted sum of the personal utilities of the individuals involved. It therefore takes a form identical to that proposed by Kitcher (*EP*, 24).

---

tions in human and ape societies that I entirely endorse. A major drawback in game theory is the inadequacy of its theories of coalition formation. For this reason, I only explore the mathematical implications of my approach in two-player games.

[7] In a (Pareto) efficient outcome, nobody can be made better off without somebody else being made worse off.

[8] There is room for confusion between the two concepts because when the luminaries of the Scottish enlightenment, notably Adam Smith (1759), first drew attention to the importance of the two notions for human morality, only the word sympathy was available, and his definition of sympathy is more closely aligned with the modern notion of empathy than with the modern notion of sympathy.

[9] Harsanyi follows Suppes in calling them extended sympathy preferences.

Empathetic preferences can therefore be understood in terms of people treating others as blood relations of various degrees of sanguinity when making comparisons of their welfare (because Kitcher's form also corresponds to Hamilton's rule). Their evolution then becomes easy to appreciate. However, it is as well to repeat that expressing an empathetic preference does not imply any personal concern for the welfare of that person.

4. The Blindfold

I think that social norms or conventions evolved as equilibrium selection devices for use in coordination games, which always have many efficient equilibria. For example, the convention of driving on the right picks out one of the two efficient Nash equilibria of the Driving Game that we play every morning on our way to work. To say so commonly provokes knee-jerk hostility, but I think that moral codes evolved for the same purpose of selecting equilibria, but in much larger games. In particular, fairness norms evolved for use in games that involve some element of sharing.

How do our fairness norms work? My thesis is that all the fairness norms that we actually use in daily life have a common deep structure that is captured in a stylized form by an idea that John Rawls (1972) called the *original position* in his celebrated *Theory of Justice* (Binmore 2005).[10] Rawls defends the device of the original position as an operationalization of Immanuel Kant's categorical imperative, but I think this is just window-dressing. The idea certainly hits the spot with most people when they hear it for the first time, but I do not believe this is because they have a natural bent for metaphysics. I think it is because they recognize a principle that matches up with the fairness norms that they actually use every day in solving the equilibrium selection problem in the myriads of small coordination games of which daily life largely consists.

Kitcher does not discuss fairness at all, but I have introduced the subject here to draw attention to the need for Adam and Eve to deploy *empathetic* preferences when behind Rawls' veil of ignorance. Recall that Eve must then bargain as though she might end up as either Adam or Eve in whatever social contract they finally agree. She must therefore assess how it would be to be Adam in one social situation as compared with Eve in another. And for what other purpose do we need the empathetic preferences that everybody agrees we are able to express?

5. Conclusion

My speculations about the manner in which our moral codes evolved differ from those of Kitcher on a wide range of issues. If space allowed, I would also have

[10] Anthropologists report that some version of the golden rule is recognized in all societies The original position is a version of the golden rule that meets the objection: Don't do unto others as they do unto you; they may not have the same tastes as yours.

liked to go on at length about our differing attitudes to the biological and social foundations of our moral institutions.[11] However, such differences are much less important than our shared view that morality can be explained without the need to look beyond the natural world.

As Kitcher observes, in defending moral naturalism, it is necessary to be able to tell at least one step-by-step story about how the moral norms that govern our behavior to some extent in modern societies might have evolved along with the societies themselves. But many such stories are possible, and ancestral social contracts leave no fossils to settle disputes over whose story is more plausible. It may even be true, as with the various hypotheses about how the eye may have evolved, that all the stories are true to some extent. What really matters at this stage is whether the stories that Kitcher tells are likely to hit the spot with traditional moral philosophers. On the whole, I think that he stands a better chance than some of his more aggressive predecessors.

Bibliography

Alger, I./J. Weibull (2012), Homo Moralis: Preference Evolution under Incomplete Information and Assortative Matching, working paper, Stockholm School of Economics

Alcock, J. (2001), *The Triumph of Sociobiology*, New York

Axelrod, R. (1984), *The Evolution of Cooperation*, New York

Binmore, K. (1994), *Playing Fair: Game Theory and the Social Contract I*, Cambridge/MA

— (1996), *Just Playing: Game Theory and the Social Contract II*, Cambridge/MA

— (2001), A Review of Robert Axelrod's *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*, URL: http://jasss.soc.surrey.ac.uk/1/1/review1.html

— (2005), *Natural Justice*, New York

— (2007), *A Very Short Introduction to Game Theory*, Oxford

— /A. Shaked (2010), Experimental Economics: Where Next?, in: *Journal of Economic Behavior and Organization* 73, 87–100

Dawkins, R. (1976), *The Selfish Gene*, Oxford

Harsanyi, J. (1977), *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge

Hofbauer, J./K. Sigmund (1998), *Evolutionary Games and Population Dynamics*, Cambridge

Hume, D. (1978[1739]), *A Treatise of Human Nature*, 2nd edition, edited by L. A. Selby-Bigge, revised by P. Nidditch, Oxford

— (1975[1777]), *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, 3rd edition, edited by L. A. Selby-Bigge, revised by P. Nidditch, Oxford

Mackie, J. (1977), *Ethics, Inventing Right and Wrong*, London

Nash, J. (1951), Non-Cooperative Games, in: *Annals of Mathematics* 54, 286–295

Nozick, R. (2001), *Invariances: The Structure of the Objective World*, Cambridge/MA

Rawls, J. (1972), *A Theory of Justice*, Oxford

[11] Kitcher emphasizes the social element, whereas I think the establishment of the deep structure of our fairness norms must have predated even language.

Skyrms, B. (1996), *Evolution of the Social Contract*, Cambridge
— (2003), *The Stag Hunt and the Evolution of the Social Structure*, Cambridge
Singer, P. (1980), *The Expanding Circle: Ethics and Sociobiology*, New York
Smith, A. (1975[1759]), *The Theory of Moral Sentiments*, edited by D. Raphael and A. Macfie, Oxford
Smith, M.-J. (1982), *Evolution and the Theory of Games*, Cambridge
Sugden, R. (1986), *The Economics of Rights, Cooperation and Welfare*, Oxford
Suppes, P. (1966), Some Formal Models of Grading Principles, in: *Synthese* 6, 284–306
Trivers, R. (1971), The Evolution of Reciprocal Altruism, in: *Quarterly Review of Biology* 46, 35–56