
F

Fostering Values of Fairness

Tadeg Quillien
University of California Santa Barbara,
Santa Barbara, CA, USA

Synonyms

[Equity](#); [Justice](#)

Definition

The processes by which evolution can design cognitive mechanisms equipped with certain assumptions about what counts as fair.

Introduction

Because natural selection designs phenotypes that favor the replication of the genes that built them, contemporary evolutionary theory predicts that organisms should be strongly biased toward their and their kin's own interest. However, in some species, including humans, the complexity of social life has led to the emergence of values that are not exclusively aligned with self-interest.

It is widely agreed that among these values is a concern for fairness. But what counts as fairness to the human mind? Depending on circumstances, it can refer to a number of different things. As an

example, what is a fair division of the spoils of a hunt among the members of the hunting party? Several criteria are possible, some of which might conflict with each other: one is equal division between the members, another is a division proportional to each party's contribution to the hunt, yet another is a division that is proportional to their need. "Fairness" has also been used to describe the impartiality of a procedure (regardless of its eventual outcome) or even a general concern for others.

Norms of fairness and their influence on human behavior have been extensively studied by psychologists and economists. Because of their diversity, they raise several issues that evolutionary theorists have recently focused on.

Explaining Deviations from Selfishness

Human behavior in a number of economic games deviates from the standard economic model, which sees agents as rational and self-interested. For instance, in a task called the dictator game, players are given a sum of money and may choose to share part of it with another player; their decision is anonymous. In contrast to the self-interested model, a substantial portion of players choose to share some of the money. This and other "anomalous" results have been taken to show that humans are averse to inequality (Fehr and Schmidt 1999). Furthermore, developmental work suggests that humans develop a sensitivity to distributive

justice from an early age: as an example, Geraci and Surian (2011) have shown that 16-month-old infants prefer looking at events in which resources are divided equally between recipients. A concern with equality raises the question of why people do not care solely about their own interests instead.

One explanation of these moral concerns is that humans are not in fact designed to look for their own interest. In models of cultural group selection, coevolution between genes and culture give rise to institutions that promote behavior for the good of the group and curtail selfishness: groups in which such institutions are prevalent outcompete other groups, causing the propagation of the institutions (Henrich et al. 2010). Although group selection models can explain generosity, they have trouble explaining why individuals would be opposed to inequality, a concern that sometimes leads to nonoptimal allocation, or even destruction, of resources and hence runs against the greater good (Shaw 2013).

Hence, most evolutionary theories explain instances of human selflessness by showing that they ultimately lead to individual benefit in the long run. These attempts focus on mutualism: in some situations, it is beneficial for both parties to interact with each other. For instance, parties in a trade benefit from the interaction, because otherwise they wouldn't take part in the exchange. In such cases, it is in the interest of an individual A to manipulate a current or potential interaction partner B so that B takes part in the interaction; often the best way to do this is to ensure that the interaction is beneficial to B. In Trivers (1971) model of reciprocal altruism, individuals are ready to suffer a cost to help another organism, if by doing so they increase the likelihood that the latter will help them in return, in a logic that can informally be described as "scratch my back and I'll scratch yours."

Recent work has extended this framework to take into account the possibility of partner choice, where people compete with each other to attract interaction partners. In the ultimatum game, a proposer offers to divide an amount of money between a responder and himself. If the responder agrees to the division, it is paid out, but if he refuses, both players get nothing. While standard game theory predicts that the proposer offers the

minimal possible nonzero amount to the responder and the latter accepts, André and Baumard (2011) show that the outcome is different when players are allowed to choose who they want to play with. In versions of their model in which it is easy to choose one's partner, evolution favors proposers who divide the resource almost equally between the responder and themselves, because by doing so they attract other players and get to take part in a greater number of beneficial interactions. When partner choice is harder, proposers can afford to make lower offers. Under this framework, what constitutes a "fair" offer is thus determined by the parties' relative bargaining power. Related game-theoretic considerations suggest that evolution should favor individuals who distribute a resource in proportion to each individual's contribution to its production, because productive individuals are more valuable partners and have better outside options. Indeed children as young as 3 years old display a tendency to distribute goods according to individual contributions (Baumard et al. 2012).

As noted above, an everyday-life example of mutually beneficial interaction is economic exchange. Behavioral economists have collected ample evidence that people have strong concerns for fairness in this domain. Kahneman et al. (1986) have shown that people object to increases in prices if they are not motivated by similar increases in production costs: it is unfair for a store owner to increase the price of snow shovels to take advantage of the increased demand caused by a snowstorm. In line with the idea that fairness norms are shaped by the parties' relative bargaining power, Friedman (2004) has proposed that this phenomenon is caused by a fundamental asymmetry of information in the relationship between buyers and sellers: while the buyer knows his reservation price (the maximum price he is willing to pay for the good), it can only be guessed by the seller. Therefore, the buyer can pretend that the current price of a good is very close to his reservation price and credibly threaten that he will not buy in case of a price increase. Because trade of goods or favors was common in man's ancestral environment (e.g., food vs. help in intragroup conflict), such commitment strategies might have evolved in this context.

The Puzzle of Impartiality

Mutualistic accounts explain why people hold values that go beyond their own self-interest, but they do not explain why fairness is often characterized by impartiality. If people preferentially care about the welfare of individuals they can have beneficial interactions with, then they should exhibit favoritism toward them. By contrast, Shaw (2013) reviews data showing that strong impartiality is an important part of human values. People sometimes value impartial individuals over individuals who treat them preferentially: for instance, children preferred a distributor who distributed four erasers equally between the subject and another child over a distributor who gave all four items to the subject. Accordingly, people strive to be impartial, even if achieving that goal means discarding resources: people imagining themselves in the role of an employer refused to give a pay raise to an employee if it was not accompanied by a raise to his equally hardworking colleague. Shaw conjectures that selection pressures related to coalitional dynamics gave rise to this concern for impartiality. In a social species, the existence of alliances might have a negative impact on outsiders, giving the latter incentives to curtail the formation of new coalitions. Punishing partiality is one way to achieve this goal, and the potential for such punishment might then encourage people to appear impartial. That people are motivated to appear impartial, rather than intrinsically valuing impartiality, is supported by experiments showing that they are less concerned with fairness when their behavior is not made public (Shaw 2013).

More generally, in an interaction between two individuals, A and B, A's interaction with another individual C matters to B, because it contains useful information. In the example of an economic transaction, in which B buys a good from A, B does not have full information about A's reservation price, but observing A selling the same good at a lower price to C gives her information about A's reservation price (namely, that it is lower than what A currently charges B), allowing

her to bargain for a better deal. Indeed, people react with outrage when they learn that someone else pays a lower price for the same good or service (Kahneman et al. 1986). Therefore, impartiality can be a means of concealing strategic information. This has broad consequences, because such strategic information can also be exploited by third parties (agents exterior to the transaction), who might be motivated to punish an ungenerous individual in order to signal that they expect better treatment from him if they happen to interact in the future, or to avoid him altogether (Krasnow et al. 2012).

Conclusion

Empirical data shows that humans hold certain early-developing, complexly patterned moral values, which point to the existence of a biological sense of fairness. There is no clear consensus yet about the exact way by which natural selection has implemented these values in us, but an emerging body of theoretical work has tackled the challenge. Fairness cannot simply be equated with generosity, and the complexity of human intuitions about what constitutes fair behavior provides a fascinating benchmark against which to compare evolutionary models.

As theories of natural selection acting at the individual level suggest (e.g., Trivers 1971), moral values might not have evolved for the good of the group. Although a concern for fairness can mitigate the negative impacts of self-interested behavior on the greater good, it also has the potential of canceling some its positive impact: for instance, by disrupting the normal operation of supply and demand, it can cause markets to fail to clear, resulting in a nonoptimal distribution of resources (Kahneman et al. 1986). Whether and how we as a society choose to foster the taste for fairness that nature has endowed us with depend on what we ultimately consider to be desirable, and careful considerations of the consequences of our intuitive preferences.

Cross-References

- ▶ [Evolution of Cooperation](#)
- ▶ [Fairness in Primates](#)
- ▶ [Game Theory](#)

References

- André, J. B., & Baumard, N. (2011). Social opportunities and the evolution of fairness. *Journal of Theoretical Biology*, *289*, 128–135.
- Baumard, N., Mascaro, O., & Chevallier, C. (2012). Preschoolers are able to take merit into account when distributing goods. *Developmental Psychology*, *48*(2), 492.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*, 817–868.
- Friedman, D. (2004). Economics and evolutionary psychology. In R. Koppl (Ed.), *Evolutionary psychology and economic theory* (pp. 17–33). Emerald: Bingley, England.
- Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science*, *14*(5), 1012–1020.
- Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., . . . Lesorogol, C. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, *327*(5972), 1480–1484.
- Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American Economic Review*, *76*, 728–741.
- Krasnow, M. M., Cosmides, L., Pedersen, E. J., & Tooby, J. (2012). What are punishment and reputation for? *PLoS One*, *7*(9), e45662.
- Shaw, A. (2013). Beyond “to share or not to share”: The impartiality account of fairness. *Current Directions in Psychological Science*, *22*(5), 413–417.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–57.