

Defection

Synonyms

Non-cooperative, cheating, selfish, exploitative, free-riding

Definition

An action that can increase an individual's payoff but that results in a socially inefficient outcome.

Introduction

In social interactions there is often a conflict between what is immediately good for the individual as opposed to what is good for the group of interacting individuals as a whole. Such situations are known as social dilemma (Dawes, 1980), and arise because actions that immediately benefit one individual may nevertheless lead to inefficient outcomes. Consider trade or exchange, for example. Cooperation would involve fairly representing the goods being offered, and honoring the terms of the exchange by, for example, not simply taking the goods of the other party by force and giving nothing in return. Defection, on the other hand, would involve misrepresenting the goods being offered or taking the goods of the other party by force. Although the total payoff to the pair of traders is maximized by cooperating, either individual can immediately gain by defecting. This occurs even though defection leads

to an inefficient outcome in which each individual is worse off than if they had cooperated.

But yet no society would function if individuals always defected and did not cooperate. This problem occurs across biological taxa, from microorganisms through to social insects and humans. A vast body of research in anthropology, computer science, economics, evolutionary biology and sociology, as well as evolutionary and social psychology, is concerned with finding conditions that take away the benefits of defection and promote cooperation.

Formalizing defection using game theory

The outcomes of defection for the actor and for the group are commonly formalized using game theory, which uses mathematical models to analyze the outcomes of strategic interactions between individuals. Table 1 shows the payoff matrix for a single-shot two-player interaction, in which each individual may either cooperate or defect. The game is symmetric, meaning that the identity of the players can be swapped without affecting the outcome. The payoff that a player receives is a measure of reward. This reward could correspond to a psychological reward or the economic concept of utility, or to biological or cultural fitness. Individuals are therefore assumed to try to choose actions that maximize their own payoff, whether this is through learning the consequences of actions, or through genetic or cultural evolution.

In this payoff matrix R represents the reward for mutual cooperation; S represents the sucker's payoff that the actor receives when it cooperates but its partner defects; T represents the temptation to defect – the payoff the actor

receives when its partner cooperates but it defects on that cooperation; and P represents the punishment payoff for mutual defection. When $R > P$ mutual cooperation leads to a higher payoff for each player than mutual defection. If $2R > S + T$ and $R > S$ then mutual cooperation yields the highest total payoff for the pair, and is hence the most efficient outcome. Given that these inequalities hold, social dilemmas occur when: 1. Unilateral defection gives a higher payoff to the defector than mutual cooperation ($T > R$), and/or 2. Mutual defection gives a higher payoff than unilateral cooperation ($P > S$). Where only condition 1 holds, this is known as a Snowdrift or Chicken game. Where only condition 2 holds, this is known as a Stag Hunt game. Where both conditions 1 and 2 hold, the situation is a Prisoner's Dilemma. In all of these games there is a tension between defection, which can give a higher payoff to the individual, and cooperation, which gives a socially efficient state. This leads to the question of why individuals would not defect.

Table 1: The payoff matrix for two-player symmetric games. Payoffs shown are for Player 1.

	Player 2 chooses Cooperate	Player 2 chooses Defect
Player 1 chooses Cooperate	R	S
Player 1 chooses Defect	T	P

Theories for why individuals would not defect: kinship and reciprocity

Much of evolutionary psychology is concerned with the evolved psychology of human hunter-gatherers. The hunter-gatherer social environment of living in small, mobile foraging groups (Boehm, 1999) corresponds to the Environment of Evolutionary Adaptedness in which our species has spent most of its evolutionary time. In this environment individuals relied heavily on cooperation with their group mates in order to obtain food for themselves and their families. This would have resulted in a value of R in Table 1 much greater than P . Evolutionary psychology then draws upon two classes of theory for why defection would not be favored.

The first is kinship. Individuals would often interact with their extended families. In such cases, the theory of kin selection from evolutionary biology (Hamilton, 1964) tells us that a genetic predisposition to cooperate rather than defect can be favored due to the fact that relatives share genes. This means that when interactions are between relatives, a gene to cooperate helps copies of itself in other individuals and so spreads in the population. In other words, when interactions are between genetic relatives then individuals in cooperative families will enjoy the payoff R , whereas individuals in families that defect will receive the smaller payoff P . The same argument can also apply when the tendency to cooperate or defect is transmitted culturally through social learning, rather than genetically. In that case we can speak of the cultural relatedness between interacting individuals (Boyd & Richerson, 2011).

The second, and complementary, theory is reciprocity. This theory originates from classical game theory in economics. It arises from the observation that social interactions are typically not one time only, single-shot.

Rather, most interactions are in fact repeated. Formally, the game in Table 1 then represents a stage game that is repeated for a number of rounds.

Repeated games allow for conditional strategies in which individuals condition their own actions on the past actions of their partner. A strategy specifies the action that an individual will take for a given history of its partner's actions. The Folk Theorem of game theory (see for example Binmore, 2005) tells us that if the interaction is repeated for an indefinite length of time, and if individuals have sufficient knowledge of the past actions of their partners, then any strategy which gives more than the minimax payoff can be an equilibrium. The minimax payoff is the payoff that a player receives when its partner tries to minimize that payoff. In Table 1, the minimax payoff would be the payoff that an individual receives if its partner always defects. Consequently, any strategy in which individuals do not always defect results in a higher payoff to each individual. Such a strategy can therefore be an equilibrium even when individuals are concerned with only maximizing their own payoffs. The reason is that any individual who deviates from the strategy can be punished by having its payoff reduced to the minimax payoff through defection by its partners. If an individual defected then it would subsequently always receive defection from other individuals, limiting its future payoffs to P or S , which are both smaller than the payoff R from mutual cooperation. The threat of being defected against can therefore stabilize cooperation when interactions are repeated.

This theory of cooperation under repeated interactions was later popularized in evolutionary biology by Trivers (1971). The Tit-for-Tat strategy advocated by Axelrod (1984) is one example of a conditional strategy.

However, in general there will be a very large number of other strategies that give more than the minimax payoff and so which can also be equilibria.

What is needed for this result to apply in a real-world scenario is that the game is indefinitely repeated and that individuals have sufficient information about the past actions of their partners. In hunter-gatherers interactions were effectively repeated for an indefinite number of times. This is because individuals relied on interactions with their group mates to survive, and no individual could predict the time of its death. Individuals also lived in close-knit communities where they repeatedly interacted with the same individuals. This made it easy for them to obtain information about the past actions of other group members.

The effect of defection on the evolved psychology of humans

The social environment of hunter-gatherers would have selected for psychological traits to detect defectors. Individuals would be selected to have a propensity to obtain and spread information about the actions of other group members, for example through gossip. The spread of information would be facilitated by certain institutional rules that groups developed (North, 1990). An example of this is the institutional rule of the whole group discussing around a campfire each evening. Individuals would also become sensitive to what other individuals thought about themselves, i.e. their reputation.

Because the Folk Theorem demonstrates that there will be a large number of possible equilibria, individuals would also be selected to coordinate their actions to ensure that they reached an equilibrium that yielded a high payoff to them. In other words, each individual would benefit from its group

coordinating on an equilibrium in which cooperation was frequent and defection rare. This would select for individuals to create institutional rules and norms that coordinated actions onto a high payoff equilibrium. For example, by using low cost coordinated punishment against acts of defection, such as ridicule and ostracism (Boehm, 1999).

Defection in large-scale societies

The origin of agriculture led to individuals living in much larger social groups. This meant that there was a lower genetic relatedness between interacting individuals. It also meant that it became harder for individuals to obtain information about the past actions of group members. Why, then, is defection not more frequent in large-scale societies? Some scholars argue that defection would actually be individually advantageous, but that individuals do not defect because they still carry the evolved psychological traits from our past hunter-gatherer environment. Under this account, our evolved psychology is “misfiring” in the new environment of large-scale societies. This “misfire” argument was popularized by Dawkins (1976). Other scholars contend that defection is not advantageous even in large-scale societies, because groups have created institutions that still allow the conditions of cooperation under the Folk Theorem to be satisfied (e.g. Greif, 2006).

Conclusion

The possibility of defection threatens the potential gains from cooperation. The ecological and social environment of hunter-gatherers would have selected for traits that supported cooperation through kinship and

reciprocity. This would include equipping individuals with a psychology that made them sensitive to their own reputation and the reputation of others. It would also have favored traits that facilitate the spread of information about the past actions of group members, such as gossip. To what extent these traits continue to be adaptive in the new environment of large-scale societies is a current research question.

Cross-references (Evolution Of Cooperation, Tit-For-Tat Cooperation, Reciprocal Altruism And Cooperation For Mutual Benefit, Robert Axelrod's (1984) *The Evolution Of Cooperation*, Tit For Tat (Axelrod & Hamilton, 1981), Game Theory, Cooperation Varies With Genetic Relatedness, Reputation and Altruism, Hunter-Gatherer Societies as Sources of Data in Evolutionary Psychology, Environment Of Evolutionary Adaptedness)

References

- Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books, Inc.
- Binmore, K. (2005). *Natural Justice*. Oxford: Oxford University Press.
- Boehm, C. (1999). *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*. Cambridge, MA: Harvard University Press.
- Boyd, R., Richerson, P., & Henrich, J. (2011). Rapid cultural adaptation can facilitate the evolution of large-scale cooperation. *Behavioral Ecology and Sociobiology*, 65, 431-444.

- Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, 31, 169-193.
- Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.
- Greif, A. (2006). *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge, UK: Cambridge University Press.
- Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7, 1-16.
- North, D. C. (1990). *Institutions, Institutional Change and Economic Performance (Political Economy of Institutions and Decisions)*. Cambridge, UK: Cambridge University Press.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46, 35-57.