Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 2686–2689, April 1996
Evolution

Human cooperation in the simultaneous and the alternating Prisoner's Dilemma: Pavlov versus Generous Tit-for-Tat

CLAUS WEDEKIND AND MANFRED MILINSKI

Abteilung Verhaltensökologie, Zoologisches Institut, Universität Bern, CH-3032 Hinterkappelen, Switzerland

Communicated by Robert M. May, University of Oxford, Oxford, United Kingdom, November 9, 1995 (received for review August 7, 1995)

ABSTRACT The iterated Prisoner's Dilemma has become the paradigm for the evolution of cooperation among egoists. Since Axelrod's classic computer tournaments and Nowak and Sigmund's extensive simulations of evolution, we know that natural selection can favor cooperative strategies in the Prisoner's Dilemma. According to recent developments of theory the last champion strategy of "win–stay, lose–shift" ("Pavlov") is the winner only if the players act simultaneously. In the more natural situation of players alternating the roles of donor and recipient a strategy of "Generous Tit-for-Tat" wins computer simulations of short-term memory strategies. We show here by experiments with humans that cooperation dominated in both the simultaneous and the alternating Prisoner's Dilemma. Subjects were consistent in their strategies: 30% adopted a Generous Tit-for-Tat-like strategy, whereas 70% used a Pavlovian strategy in both the alternating and the simultaneous game. As predicted for unconditional strategies, Pavlovian players appeared to be more successful in the simultaneous game whereas Generous Tit-for-Tat-like players achieved higher payoffs in the alternating game. However, the Pavlovian players were smarter than predicted: they suffered less from defectors and exploited cooperators more readily. Humans appear to cooperate either with a Generous Tit-for-Tat-like strategy or with a strategy that appreciates Pavlov's advantages but minimizes its handicaps.

In the Prisoner's Dilemma the players can either cooperate or defect (not cooperate). If they cooperate, both do better than if they had both defected. But if one player defects while the other cooperates, the defector gets the highest reward and the cooperator gets the lowest reward. Each should defect no matter what his opponent does if they meet only once. So both end up with a much lower reward than they could have gained if they had decided to cooperate. Hence the dilemma. If the game is played repeatedly by the same players cooperation by reciprocal altruism (1) is possible (2–7).

In two computer tournaments for which Axelrod (3) solicited deterministic strategies from game theorists, a very simple cooperative strategy called Tit-for-Tat (TFT) was the winner. TFT starts cooperatively and then repeats the opponent's previous move. A weakness of deterministic TFT shows up when players can make mistakes: they are caught in mutual retaliation until the next mistake occurs. In computer simulations of evolution with randomly generated mixtures of stochastic strategies that respond to the opponent's last move, Nowak and Sigmund (4) found that a "Generous TFT" was the evolutionary end product. Generous TFT "corrects" mistakes by being cooperative with a certain probability after the partner's defection. When, in further computer simulations, a strategy could react not only to the partner's but also to its own previous move and any mutant strategy was allowed for (5) a new champion appeared in >80% of the simulations: "Win–stay, lose–shift" [or "Pavlov" (8, 9)]. It cooperates after both

		Opponent C	Opponent D
Player	C	3	0
	D	4	1

FIG. 1. Payoff matrix of the game showing the payoffs to the player. Each player can either cooperate (C) or defect (D). In the simultaneous game this payoff matrix could be used directly; in the alternating game it had to be translated: if a player plays C, she gets 4 and the opponent 3 points. If she plays D, she gets 5 points and the opponent 0. Two such subsequent choices are equivalent (plus a constant) to one choice in the simultaneous game (7).

partners have cooperated with probability $P_1 = 1$, after self cooperated and partner defected with $P_2 = 0$, after self defected and partner cooperated with $P_3 = 0$, and after both defected with $P_4 =$ (almost) 1. The greatest difference between Pavlov and Generous TFT appears in P_3. Pavlov cooperates with TFT and itself, exploits unconditional cooperators, but it is more heavily exploited by unconditional defectors than TFT. Although there is supporting evidence for TFT-like cooperation from experiments with animals (10–15), no experimental study has tested specifically for Generous TFT or for Pavlov.

Models of the evolution of cooperation assumed that the two players make their decisions in synchrony, which seems improbable in many biological situations (6, 7). Although the iterated Prisoner's Dilemma is a suitable model for both simultaneous and alternating choices (16), the most frequent evolutionary outcome has been shown to differ in these situations (6, 7). A Pavlov strategy which does so well in the simultaneous game (5) does rather poorly in the strictly alternating case (6, 7). A kind of Generous TFT [also called "Firm but Fair" (6)] is the predicted evolutionary end product in the alternating game (for the payoff matrix shown in Fig. 1, the winning strategy is described by $P_1 = 1, P_2 = 0, P_3 = 1, P_4 = 2/3$). In the computer simulations, only a memory of one move of both players was allowed for.

We did two experiments with first-year biology students at Bern University. Humans have been and still are under both simultaneous and alternating Prisoner's Dilemma selection (17) and could thus have evolved or learned suitable strategies. They have been shown to be able to cooperate in the Prisoner's Dilemma game (18, 19). The aim of this study was to test specifically whether the strategy that the students use in a simultaneous or an alternating Prisoner's Dilemma is closer to Pavlov or to the Generous TFT described above. A secondary aim was to describe the students' strategies quantitatively so that our results can be used to test for any other strategy that might be predicted in the future.

METHODS

The students were assigned to four groups (60% women, similar percentage in each group, $\chi^2 = 2.26$, df = 3, $P = 0.52$).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: TFT, Tit-for-Tat.

Table 1. Nested ANOVA (two groups of players per simultaneous and alternating conditions) with four replicates on payoff per game, and proportion of cooperative moves per game in the first session

Source of variation	SS	df	F	P
Payoff per game				
Between subjects				
Mode*	0.004	1	0.0007	0.98
Group	7.44	2	6.58	0.003
Error	30.49	54		
Within subjects (four games played)				
Payoff	11.22	3	6.47	<0.001
Payoff × mode*	5.02	3	2.89	0.037
Payoff × group	8.25	6	2.35	0.034
Error	93.72	162		
Proportion of cooperative moves per game				
Between subjects				
Mode*	0.0002	1	0.0018	0.97
Group	1.84	2	8.11	<0.001
Error	6.12	54		
Within subjects (four games played)				
C proportion†	2.81	3	11.13	<0.001
C proportion × mode*	1.25	3	4.97	0.003
C proportion × group	2.01	6	3.99	<0.001
Error	13.63	162		

Both payoff and proportion of cooperative moves increase from the first to the last game.
*Mode, Prisoner's Dilemma with either simultaneous or alternating moves.
†C proportion, proportion of cooperative moves per game.

They had not heard about the Prisoner's Dilemma in their courses yet. The groups were tested sequentially on the same day. Groups 1 (n = 16) and 3 (n = 14) played the alternating and groups 2 (n = 14) and 4 (n = 14) played the simultaneous Prisoner's Dilemma in both their first and second session. Each group was instructed in the same way (by M.M.; the text of the instructions is available from the authors on request) about the rules and the prices for the three highest-ranking players in each group. To provide the students with a social situation for learning the game in a first session, each subject had to play four games against randomly assigned members of the group. The rest of the group sat in the back of the room watching the players and a large screen in front of the players. Only the last pair of choices and both payoffs (Fig. 1) were displayed to facilitate a short memory. The players were separated by an opaque partition. Each player indicated her choice by a card with either a C or a D. The operator (C.W.) supplied a computer program with the choices which displayed each choice and the resulting payoffs separately in the alternating game or simultaneously in the simultaneous game, respectively. The program determined on line randomly (*i*) each pair of players under the constraint that each student had to play four games each with a different partner, (*ii*) the player who had to start in the alternating game, and (*iii*) the end of a game: an even number drawn from a uniform distribution with a tail (three overlaid normal distributions with means of 1, 17, and 33, respectively, and a SD of 6.7 each); under the constraint that each player had at least two choices, a player had actually up to 24 choices. After each game the mean payoff per choice of both players was displayed as their payoffs from this game. The subjects were instructed that, in each group, the players with the three highest mean payoffs after the second session would receive 60 Swiss francs (SFr), 30 SFr, and 10 SFr, respectively. So the subjects knew that it was important to gain as many points as possible rather than to beat the other player.

In the second session each student had to play one game against each of two nonmembers of the group without audience. Otherwise, the procedure was analogous to the first

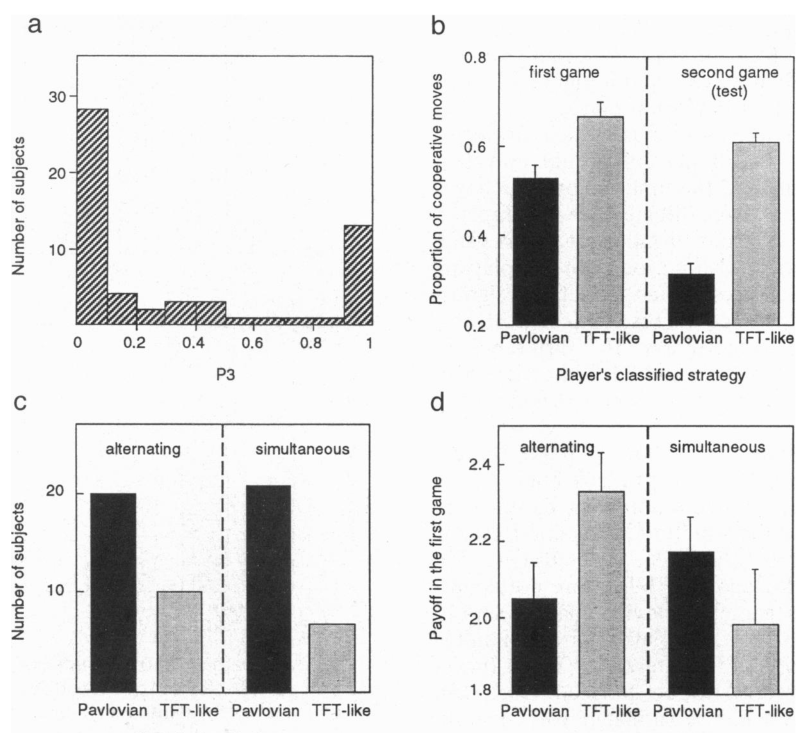

FIG. 2. Distribution and success of strategies classified as either Pavlovian or TFT-like. (*a*) Distribution of P_3 values of all players in the second game against the all-C pseudoplayer. (*b*) TFT-like players had a higher proportion (mean + SE) of cooperative choices than Pavlovian players in both the first session (t = 2.67, df = 56, P = 0.01, two-tailed) and the second session (t = 7.31, df = 56, P < 0.0001, two-tailed). (*c*) Percentage of Pavlovian and TFT-like players in the first session under both alternating and simultaneous conditions. (*d*) Mean + SE payoff of Pavlovian and TFT-like players in the alternating and the simultaneous first session.

session. The game was terminated after 20 choices. These two pseudoplayers, a male and a female student, played predetermined strategies designed to test for Pavlov and Generous TFT, respectively—i.e. to detect the P_1, P_2, P_3, and P_4 values of the subject's strategy. The pseudoplayers always started in the alternating game and either (*a*) played always C with a single D in their 16th choice ("all C") or (*b*) made their first five choices according to strict TFT and subsequently played D ("all D"). The pseudoplayers' sequence and strategy were both alternated between subjects. One pair of pseudoplayers was used for groups 1 and 2, a different pair for groups 3 and 4. After the second session the subjects had no contact with group members that waited for their game watching a video on birds-of-paradise.

RESULTS

During the four games of the first session the subjects became more and more cooperative and successful (Table 1). They achieved almost the same payoff and readiness to cooperate in the alternating as in the simultaneous Prisoner's Dilemma. However, the development (Table 1) and the pattern of the play differed between the two modes (discriminant analysis with the average frequency of eight possible combinations of three sequential moves, e.g., C C D; $r^2 = 0.97$, $F_8 = 198.2$; $P < 0.0001$). Thus, the alternating and the simultaneous Prisoner's Dilemma were different games, as expected (6, 7).

Because the greatest difference between Pavlov and Generous TFT is predicted for P_3 (probability to play C after self's D and partner's C), we determined the P_3 value of each subject's response in the second session against the all-C pseudoplayer (Fig. 1). The distribution of P_3 values across subjects appeared to be bimodal (Fig. 2*a*) with peaks at 0 and 1, as expected for Pavlov and any TFT, respectively. We further classified the response of each subject to our all-C pseudoplayer into either Pavlovian or TFT-like, depending on which type of strategy (Pavlov or TFT) would correspond to the player's response with fewer mistakes. When necessary (10 cases) we used the P_3 value for a decision. It turned out that each of the classified TFT-like players had a higher P_3 value than each subject classified as a Pavlovian player.

If our Pavlovian and TFT-like players used their strategies consistently, the P_3 values of all players should correlate positively between the second and the first session, which was the case [$r = 0.24$, $n = 54$, $p = 0.05$, directed; we use directed instead of one-tailed tests to avoid inflation of the α value (20)]. Such a consistency could also be found in the players' proportion of cooperative choices, which correlated significantly between the first and the second session ($r = 0.50$, $n = 58$, $P < 0.0001$, directed). Accordingly, the difference in cooperativity between our Pavlovian and TFT-like players in the first session was similar to that in the second session (Fig. 2*b*).

There were more Pavlovian (70%) than TFT-like players overall ($\chi^2 = 10.24$, df = 1, $P = 0.002$, two-tailed), but contrary to our expectation their proportion did not differ significantly between the alternating and the simultaneous game (test for heterogeneity, $\chi^2 = 0.49$, df = 1, $P = 0.31$, directed) (Fig. 2*c*). If the strategies, either Pavlovian or TFT-like, are not conditional on the mode of the game, Pavlovian players should achieve higher payoffs than TFT-like players in the simultaneous game and lower payoffs in the alternating mode (6, 7), an expectation which appeared to be fulfilled (Fig. 2*d*). Although the two strategies achieved similar payoffs overall, which of the two strategies was more successful was dependent of the mode (alternating or simultaneous) (two-way ANOVA on payoff: for effect of mode; $F = 0.90$, df = 1, n.s.; for effect of player's strategy; $F = 0.15$, df = 1, n.s.; for interaction between mode and strategy: $F = 3.90$, df = 1, $P = 0.03$, directed). Neither the subjects' nor the pseudoplayers' sex had any significant effect. No significant group effect on the proportion of each strategy could be found ($\chi^2 = 5.5$, df = 3, $P = 0.154$, two-tailed).

Although the P_3 value is the most important for distinguishing TFT-like strategies from Pavlovian strategies, we can also determine P_1, P_2, and P_4 in the second session for all players to describe their strategies more precisely. As was the case with P_3, the players were consistent in the other P values between the first and the second session (for P_1, $r = 0.54$, $P < 0.0001$; for P_2, $r = 0.31$, $P = 0.01$; for P_4, $r = 0.27$, $P = 0.02$, directed). The P_1 and P_3 values of our TFT-like players were very close to those predicted for Generous TFT (Fig. 3*a*). They were less cooperative than expected in P_4 and more cooperative in P_2. Thereby, they achieved a higher payoff against both our all-C pseudoplayer (Wilcoxon, $P = 0.045$, two-tailed) and our all-D pseudoplayer ($P = 0.02$) (Fig. 3*b*). Since P_4 was clearly >0, the TFT-like players used a kind of Generous TFT rather than TFT. Our Pavlovian players were more cooperative than expected in their P_2 and P_3, and less cooperative than expected in their P_1 and P_4. Especially their P_4 was much lower than expected (Fig. 3*a*) and on average 0.10 smaller than their P_4 in the first session. By their deviation from a more strict Pavlovian strategy they achieved significantly higher payoffs against both the all-C ($P < 0.001$) and the all-D ($P < 0.0001$) pseudoplayer than a more classic Pavlovian strategy would have gained (Fig. 3*b*).

DISCUSSION

The subjects used rather consistently either a Generous TFT-like or a Pavlovian strategy. We regard it as Pavlov-like because it has, as expected for a Pavlovian strategy, a high P_1 and low P_2 and P_3, but contrary to classic Pavlov it is smarter

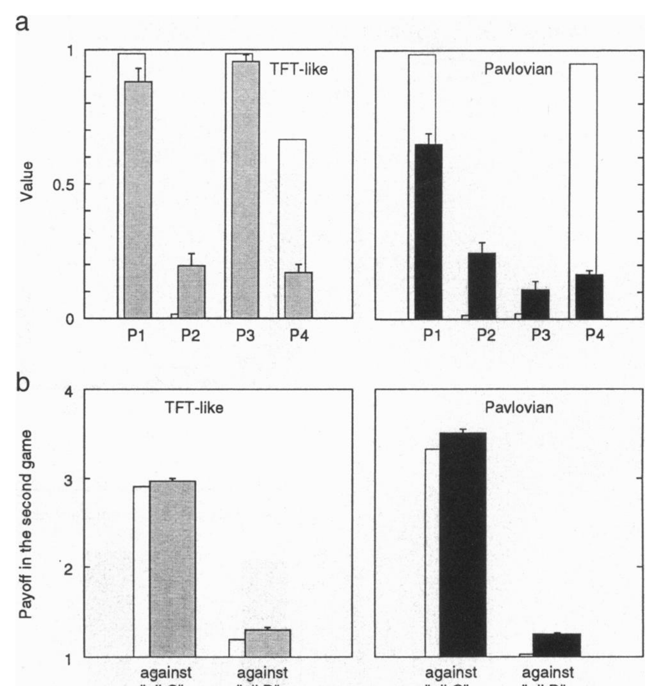

FIG. 3. Comparison of the subjects' strategies with either Generous TFT or pure Pavlov in the second session. (*a*) Mean + SE P_1, P_2, P_3, and P_4 values of classified TFT-like and Pavlovian players. Expected values (open bars) for Generous TFT with our payoff matrix and memory of one step (7) are 1, 0, 1, and 2/3; expected values for Pavlov are: 1, 0, 0, and 0.95; all values of 0 and 1 are depicted as being slightly less extreme to allow for stochasticity (5). (*b*) Mean + SE payoff of the subjects' strategies classified as either TFT-like or Pavlovian compared with the expected strategies (see above) (7) in both the alternating and the simultaneous game.

when playing against unconditional defectors: by its low P_4 it is much less exploited than the classic strategy. This has been predicted for an improved Pavlov (9). Our Pavlovian players were even better than that: by their P_1 value that was somewhat lower than expected they exploited unconditional cooperators sooner than classic Pavlov would do. It looks as though humans use strategies in the iterated Prisoner's Dilemma that are more sophisticated than were expected from computer simulations with the unavoidably restrictive assumption of a short memory (6, 7). Although we displayed only the last pair of choices on the screen, it is possible that the subjects achieved higher payoffs by using a longer memory. There are no predictions yet from theoretical investigations for the effect of a longer memory on the strategies to be used. Both theoretical and experimental studies of this relation are badly needed.

Because we found both a Generous TFT-like and a Pavlov-like strategy in our students, both the simultaneous and the alternating Prisoner's Dilemma could have been part of human ecology. The missing flexibility of our subjects to play Pavlov in the simultaneous and Generous TFT in the alternating game as predicted (6, 7) suggests either that people might have preferred niches, which contain either the simultaneous or the alternating situation, or that our game situations did not offer the natural clues by which humans would recognize the mode and trigger their response conditionally. Nevertheless, the subjects' play differed in the two modes and, as predicted (6, 7), Pavlovian players appeared to gain more in the simultaneous game whereas Generous TFT-like players were better off in the alternating mode. To understand the coexistence of our two consistent types of players it may be helpful to find out more about their other social strategies, including their use of memory of past social interactions.

We thank the students for their participation; Tatjana Hadermann, Maddalena Tognola, Patrick Boltshauser, and Reto Künzler for assisting as pseudoplayers; and Markus Frischknecht, Alex Kacelnik, Anne Magurran, Martin Nowak, and Karl Sigmund for comments on the manuscript.

1. Trivers, R. L. (1971) *Q. Rev. Biol.* **46,** 35–57.
2. Axelrod, R. & Hamilton, W. D. (1981) *Science* **211,** 1390–1396.
3. Axelrod, R. (1984) *The Evolution of Cooperation* (Basic Books, New York).
4. Nowak, M. A. & Sigmund, K. (1992) *Nature (London)* **355,** 250–253.
5. Nowak, M. A. & Sigmund, K. (1993) *Nature (London)* **364,** 56–58.
6. Frean, M. R. (1994) *Proc. R. Soc. London B* **257,** 75–79.
7. Nowak, M. A. & Sigmund, K. (1994) *J. Theor. Biol.* **168,** 219–226.
8. Kraines, D. & Kraines, V. (1988) *Theory Decision* **26,** 47–79.
9. Kraines, D. & Kraines, V. (1993) *Theory Decision* **35,** 107–150.
10. Packer, C. (1977) *Nature (London)* **265,** 441–443.
11. Wilkinson, G. S. (1984) *Nature (London)* **308,** 181–184.
12. Lombardo, M. P. (1985) *Science* **227,** 1363–1365.
13. Milinski, M. (1987) *Nature (London)* **325,** 433–435.
14. Dugatkin, L. A. (1991) *Behav. Ecol. Sociobiol.* **29,** 127–132.
15. Axelrod, R. & Dion, D. (1988) *Science* **242,** 1385–1390.
16. Boyd, R. (1988) *Ethol. Sociobiol.* **9,** 211–222.
17. Milinski, M. (1993) *Nature (London)* **364,** 12–13.
18. Dawes, R. M. (1980) *Annu. Rev. Psychol.* **31,** 169–193.
19. Frank, R. H., Gilovich, T. & Regan, D. T. (1993) *Ethol. Sociobiol.* **14,** 247–256.
20. Rice, W. R. & Gaines, S. D. (1994) *Trends Ecol. Evol.* **9,** 235–237.