

Working memory constrains human cooperation in the Prisoner's Dilemma

MANFRED MILINSKI* AND CLAUS WEDEKIND

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

Edited by Robert Axelrod, University of Michigan, Ann Arbor, MI, and approved September 15, 1998 (received for review June 8, 1998)

ABSTRACT Many problems in human society reflect the inability of selfish parties to cooperate. The “Iterated Prisoner’s Dilemma” has been used widely as a model for the evolution of cooperation in societies. Axelrod’s computer tournaments and the extensive simulations of evolution by Nowak and Sigmund and others have shown that natural selection can favor cooperative strategies in the Prisoner’s Dilemma. Rigorous empirical tests, however, lag behind the progress made by theorists. Clear predictions differ depending on the players’ capacity to remember previous rounds of the game. To test whether humans use the kind of cooperative strategies predicted, we asked students to play the iterated Prisoner’s Dilemma game either continuously or interrupted after each round by a secondary memory task (i.e., playing the game “Memory”) that constrained the students’ working-memory capacity. When playing without interruption, most students used “Pavlovian” strategies, as predicted, for greater memory capacity, and the rest used “generous tit-for-tat” strategies. The proportion of generous tit-for-tat strategies increased when games of Memory interfered with the subjects’ working memory, as predicted. Students who continued to use complex Pavlovian strategies were less successful in the Memory game, but more successful in the Prisoner’s Dilemma, which indicates a trade-off in memory capacity for the two tasks. Our results suggest that the set of strategies predicted by game theorists approximates human reality.

The “iterated Prisoner’s Dilemma” has become the paradigm for the evolution of the cooperation of egoists (1–4). Players in this game can either cooperate or defect (not cooperate). If they cooperate, both do better than if both defect. If one player defects while the other cooperates, the defector gets the highest reward, and the cooperator gets the lowest. A rational player should defect no matter what his opponent does if they play only one round; thus, both players will end up with a much lower reward than they would have received if they had cooperated—hence, the dilemma. If the game is played repeatedly by the same players, cooperation by reciprocity (5) is possible (1, 6–8).

In computer simulations of evolution with randomly generated mixtures of stochastic strategies that respond only to the opponent’s last move (memory-1 strategies), the “generous tit-for-tat” (GTFT) strategy was the evolutionary end product (7). GTFT players usually copy their partner’s last choice but are sometimes cooperative after their partner’s defection. When players adopt strategies that also react to their own previous move (memory-2 strategies), a new winning strategy emerged (8), the “win-stay, lose-shift” or “Pavlov” strategy. With this strategy, one cooperates after both players have cooperated with probability $P_{cc} = 1$, after one has cooperated and one’s partner has defected with $P_{cd} = 0$, after one has

defected and one’s partner has cooperated with $P_{dc} = 0$, and after both defected with $P_{dd} =$ (almost) 1. The greatest difference between Pavlovian and GTFT strategies is in P_{dc} , which is 1 for GTFT. Pavlovian players cooperate with GTFT players and other Pavlovian players, exploit unconditional cooperators, and are more heavily exploited by unconditional defectors than GTFT players. When longer memories were allowed (9–11), memory-4 strategies evolved (taking into account the previous two choices of both players) that were Pavlovian but answered a single defection by defecting twice; their P_{dd} would be lower than that of strict Pavlovian memory-2 strategies.

Humans have been and still are under Prisoner’s Dilemma selection (12) and thus could have evolved or learned suitable strategies. They are able to cooperate in the Prisoner’s Dilemma (13, 14). A rigorous test of the results of simulations of evolution (7–11) would be to constrain the subjects’ working-memory capacity experimentally; compared with controls, these subjects should use the predicted memory-1 strategy GTFT more often. In a study with unconstrained subjects, only a minority used GTFT, the rest used Pavlovian strategies (14). We further predict that memory-constrained subjects who adopt a GTFT strategy will profit from a trade-off of their memory capacity.

In our previous study (14), we found that the Pavlovian strategy actually used was both more complex (i.e., its P_{dd} was much smaller than 1) and significantly more successful than strict Pavlov. Because complex strategies are harder to remember and implement, there are costs for complexity (15). Theorists have dealt with complexity in various plausible ways (16–18). A native efficient strategy can only be invaded by strategies that are as efficient but simpler (15–18). A strategy will be displaced by another one of greater complexity if the resulting improvement in game payoffs is sufficiently large (18). Thus, if complex Pavlovian strategies become more costly when the subjects’ memory capacity is constrained, we predict, by using this theory, that either simpler Pavlovian or even less complex memory-1 strategies (i.e., the strategies that survived simulations of evolution with short strategic memories) will emerge (6–8).

Humans have multiple memory systems that include a working (short-term) memory and various long-term systems (19–25). Whereas long-term memory has an enormous capacity for storage, coupled with relatively slow input and retrieval, working memory has a limited capacity for storage with rapid input and retrieval (19, 20). Often, a single training trial produces only short-term memory (22). Unrehearsed information is forgotten within seconds or sometimes minutes (19, 22); the most recent information is forgotten last (19).

By supplying subjects with an additional memory task (i.e., playing the game “Memory”) after each choice in the Prisoner’s Dilemma, many factors will work to reduce the number of

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.

© 1998 by The National Academy of Sciences 0027-8424/98/9513755-4\$2.00/0
PNAS is available online at www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office. Abbreviations: GTFT, generous tit-for-tat; TFT, tit-for-tat.

*To whom reprint requests should be addressed. e-mail: manfred.milinski@esh.unibe.ch.

choices that will be remembered: (i) the recency effect is dissipated by the intervening task; (ii) the additional memory task uses up storage capacity; (iii) the additional memory task prevents rehearsal of previous choices and thus their transfer into long-term memory, and (iv), because the second task takes time, stored information is more likely to decay. If playing Memory reduces the subjects' working-memory capacity for the Prisoner's Dilemma, we predict that they will use either simpler Pavlovian strategies or switch to memory-1 strategies (i.e., any GTFT).

METHODS AND RESULTS

First-year biology students of Bern University were assigned either to a constrained-memory group (1996 experiment, $n = 16$; 1997 experiment, $n = 16$) or to an unconstrained-memory group (1996 experiment, $n = 14$; 1997 experiment, $n = 16$). All groups played the simultaneous Prisoner's Dilemma ($T = 4$, $R = 3$, $P = 1$, $S = 0$; ref. 14). A cooperating player receives R points from a cooperating partner but only S points from a defecting one. A defecting player gets T points from a cooperating partner but only S points from a defecting one. In the first session, to learn the game in a social situation, each subject played four games against randomly assigned members of the group. Each player chose to cooperate or defect by lifting a card labeled either "C" or "D". The most recent pair of choices and payoffs were displayed simultaneously on a large screen opposite the players and the group after both players had chosen. An on-line program randomly terminated each game (14); up to 24 choices occurred.

In the constrained-memory groups, each player uncovered two cards of her Memory game (on the table next to an opaque partition between the players) after each choice in the Prisoner's Dilemma game. The game consisted of 32 cards showing either Swiss money or blanks (8 cards). If the cards showed the same picture, they were removed; otherwise, they were returned. For Memory players, the mean payoff per choice in the Prisoner's dilemma was multiplied by the mean payoff per round of Memory, so that nobody could win by concentrating on only one task. The subjects knew that, in each group, the players with the first, second, and third highest mean payoffs after the second session would receive 60, 30, and 10 Swiss Francs, respectively.

In the second session, each student played two games of 20 choices (number not known by the student), each against a different "pseudoplayer" who was not a member of the group. Games in the second session were played without an audience. Otherwise, the procedure was analogous to the first session and included playing Memory in the constrained-memory groups. The pseudoplayers used predetermined strategies to test for Pavlov or GTFT (i.e., to determine the P_{cc} , P_{cd} , P_{dc} , and P_{dd} values of a subject's strategy). Either the pseudoplayers always chose C with the exception of a single D for their 16th choice ("allC"), or they made their first 5 choices according to strict tit-for-tat (TFT, i.e. it starts with C and thereafter copies the partner's previous choice exactly) and subsequently played D ("allD")

Do constrained-memory players change their strategy in the Prisoner's Dilemma from Pavlovian strategies in the direction of the predicted memory-1 strategy GTFT? The greatest difference between Pavlov and GTFT is predicted for P_{dc} (probability to play C after one plays D and one's opponent plays C). P_{dc} increased significantly ($t = 2.38$, $P = 0.013$, directed) in the constrained-memory players' strategies. This result suggests that more constrained- than unconstrained-memory players used a GTFT strategy. Compatible with this hypothesis is the finding that the payoff against the allC pseudoplayer was higher for unconstrained-memory players (average payoff = 3.31) than for constrained-memory players

(average payoff = 3.12; $t = 2.19$, $P = 0.02$, directed). Pavlovian strategies exploit allC, whereas GTFT does not.

To test the hypothesis further, we classified the response of each subject to our allC pseudoplayer into either Pavlovian or TFT-like depending on which type of strategy (Pavlov or TFT) corresponded to the player's response with fewer mistakes (as in ref. 14). If this did not allow for a decision, we used also the P_{dc} value from the first session (six cases). The proportion of TFT-like players was significantly higher in the constrained- than in the unconstrained-memory groups (Fig. 1A and B). Furthermore, our Pavlovian and TFT-like players used their strategies consistently; the P_{dc} values of all players correlated positively between the second and the first session ($r = 0.32$, $n = 52$, $P = 0.015$, directed). The TFT-like strategies looked similar in constrained- and unconstrained-memory players (Fig. 1D); because their P_{cc} and P_{dc} values were almost 1 and their P_{cd} and P_{dd} values were much smaller than 1 but above 0, this strategy can be identified as GTFT (7).

Do Pavlovian constrained-memory players change their strategy in the Prisoner's Dilemma in the direction of strict Pavlov (i.e., the predicted memory-2 strategy)? If so, then P_{cd} should decrease, and P_{dd} should increase in Pavlovian constrained-memory players. Only P_{cd} (probability to play C after one chooses C and one's partner chooses D) changed significantly (also after Bonferroni adjustment, $t = 2.80$, $P = 0.005$, directed), as expected (Fig. 1C). The P_{dd} remained low, which could indicate either a Pavlovian memory-4 strategy or the memory-2 strategy GRIM that appeared in simulations of evolution but was not an end product (8).

The change in the P_{dc} strategy from the unconstrained- to the constrained-memory situation allowed for more memory in the Memory game; the payoff gained in the Memory game increased with P_{dc} (Fig. 2A). Similarly, if the change from Pavlovian strategies to GTFT allows for more memory in the Memory game, GTFT players should gain more in the Memory game than Pavlovian players among the constrained-memory players, which was the case (Fig. 2B). The analogous comparisons between the payoff gained in the Prisoner's Dilemma

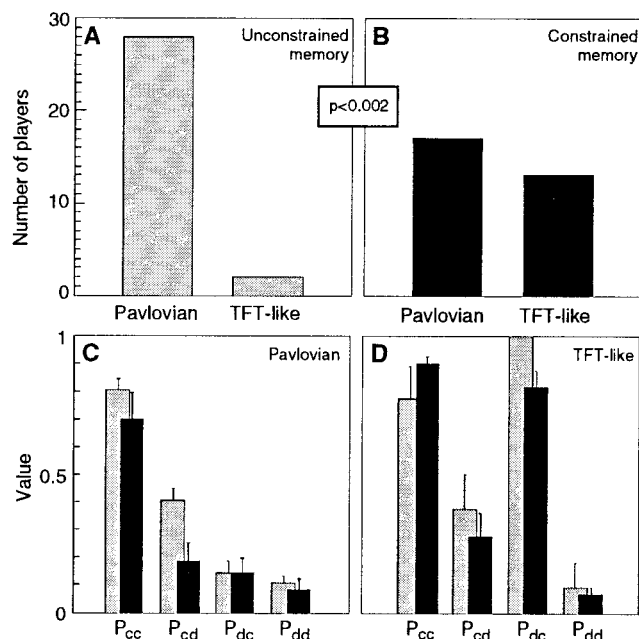


FIG. 1. Pavlovian and Tft-like players in unconstrained-memory situation (A) and constrained-memory situation (B). The difference between A and B is $\chi^2 = 8.89$, Yates corrected, directed. (C and D) Comparison of the strategy played by Pavlovian players and Tft-like players in unconstrained- (gray) and constrained-memory (black) situations.

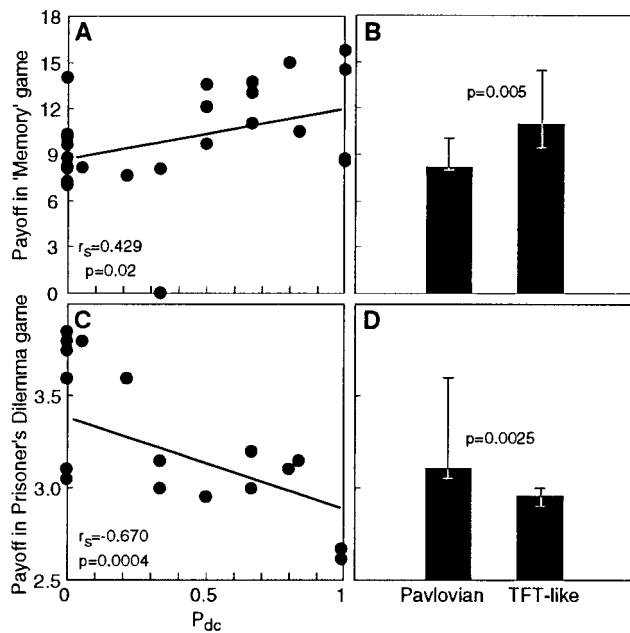


FIG. 2. The payoff achieved in the Memory game while playing against the allC pseudoplayer compared to the P_{dc} -value (A; Spearman correlation test) and to the classification into Pavlovian and TFT-like strategies (B; medians and quartiles; Mann-Whitney U test, $u = 47.5$, directed). The payoff that the constrained-memory players achieved in the Prisoner's Dilemma game while playing against the allC pseudoplayer compared to the P_{dc} -value (C; Spearman correlation test) and compared to the classification into Pavlovian and TFT-like strategies (D; medians and quartiles; Mann-Whitney U test, $u = 178.5$, directed).

game and P_{dc} (Fig. 2C) or the strategies played (Fig. 2D) indicate a trade-off in using memory capacity for the two tasks.

Our test relies on the assumption that playing Memory constrained our subjects' memory capacity so much so that many of them could remember at most the last round of choices in the Prisoner's Dilemma game. To test this assumption, two other groups of students, after a comparable first session, played the Prisoner's Dilemma game against a pseudoplayer (C, C, C, D, D, D, thereafter random). Every second subject played Memory after each choice. After the 20th choice, each subject was asked to write down both players' 20th, 19th, 18th, etc. choices from memory. Unconstrained-memory players had about 75% correct recall of both the 20th and the 19th round, whereas constrained-memory players achieved this percentage of recall for the 20th round only (Fig. 3). This result suggests that playing Memory constrained the subjects to memory-1 or memory-2 strategies in the Prisoner's Dilemma. Surprisingly, the pseudoplayer's first three choices of C were retained in some subjects' long-term memory (Fig. 3). Again, among the constrained-memory players, GTFT players gained more in the Memory game than Pavlovian players; the payoff achieved in the Memory game correlated significantly with the P_{dc} value ($r_s = 0.752$, $n = 13$, $P < 0.003$, directed), whereas the payoff achieved in the Prisoner's Dilemma game tended to decrease with increasing P_{dc} value ($r_s = -0.411$, $n = 13$, $P = 0.10$, directed).

DISCUSSION

We regard human cooperation as suitable for testing the predictions of evolutionary models. We have shown that humans adopt strategies in the iterated Prisoner's Dilemma that are conditional on their memory capacity. Memory capacity thus seems to be a constraint to social behavior. The present test situation is certainly unnatural in many respects. However, the fact that we found a significant change of

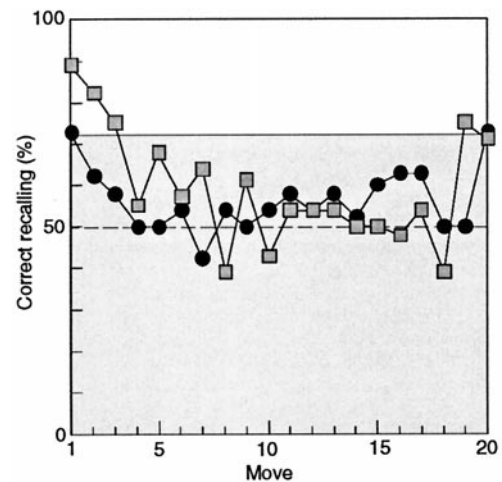


FIG. 3. Percentage of correct recall of round 20 to 1 of the Prisoner's Dilemma game by unconstrained-memory players (gray, $n = 14$) and constrained-memory players (black, $n = 13$). Correct recall = 1; incorrect = 0; undecided = 0.5. White area shows significant deviations from 50%, $P < 0.05$, one-tailed (no negative memory assumed). Correct recall on round 19 is significantly different between unconstrained- and constrained-memory players ($z = 1.871$, $P = 0.038$, directed).

strategy in the direction of the predicted (7) memory-1 strategy GTFT when we constrained the subjects' working memory capacity by a secondary memory task, is a strong indication not only of the existence of human cooperation strategies that are conditional on memory capacity but of the unexpected accuracy of the theorists' findings. Humans seem to use the simple rule of GTFT, as predicted, when they trade-off working-memory capacity. The simple TFT strategy that won Axelrod's tournaments (1, 6) was not far from human reality. Conditionally cooperative strategies such as TFT have the highest evolutionary robustness as formal models have shown (15, 26). We found either complex Pavlovian strategies or simple GTFT; no memory-constrained subject used strict Pavlovian strategies as a "stepping stone" of intermediate complexity, as might be expected. We are simpletons but only when it pays off.

We thank the students for their participation, P. Aeschlimann, P. Boltshauser, M. Christen, M. Frischknecht, R. Künzler, D. Mazzi, N. Steck, K. Turi Nagy, and M. Tognola for assisting as pseudoplayers, M. Nowak, K. Sigmund, and the anonymous reviewers for comments on the manuscript, and the Swiss National Science Foundation for support.

1. Axelrod, R. & Hamilton, W. D. (1981) *Science* **211**, 1390–1396.
2. May, R. M. (1981) *Nature (London)* **292**, 291–292.
3. May, R. M. (1987) *Nature (London)* **327**, 15–17.
4. Dugatkin, L. A. (1997) *Cooperation Among Animals: An Evolutionary Perspective* (Oxford Univ. Press, Oxford).
5. Trivers, R. L. (1971) *Q. Rev. Biol.* **46**, 35–57.
6. Axelrod, R. (1984) *The Evolution of Cooperation* (Basic Books, New York).
7. Nowak, M. A. & Sigmund, K. (1992) *Nature (London)* **355**, 250–253.
8. Nowak, M. & Sigmund, K. (1993) *Nature (London)* **364**, 56–58.
9. Axelrod, R. (1987) in *Genetic Algorithms and Simulated Annealing*, ed. Lawrence, D. (Pitman, London), pp. 32–41.
10. Lindgren, K. (1991) in *Artificial Life II*, eds. Langton, C. G., Farmer, J. D., Rasmussen, S. & Taylor, C. (Addison-Wesley, Reading, MA), pp. 295–312.
11. Hauert, C. & Schuster, H. G. (1997) *Proc. R. Soc. London Ser. B* **264**, 513–519.
12. Milinski, M. (1993) *Nature (London)* **364**, 12–13.
13. Dawes, R. M. (1980) *Ann. Rev. Psychol.* **31**, 169–193.

14. Wedekind, C. & Milinski, M. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 2686–2689.
15. Bendor, J. & Swistak, P. (1997) *American Political Science Review* **91**, 290–307.
16. Rubinstein, A. (1986) *J. Econ. Theory* **39**, 83–96.
17. Banks, J. S. & Sundaram, R. K. (1990) *Games Econ. Behav.* **2**, 97–117.
18. Binmore, K. G. & Samuelson, L. (1992) *J. Econ. Theory* **57**, 278–305.
19. Baddeley, A. (1987) *Working Memory* (Clarendon, Oxford).
20. Baddeley, A. (1994) in *Memory Systems 1994*, eds. Schacter, D. L. & Tulving, E. (MIT Press, Cambridge, MA), pp. 351–367.
21. Craik, F. I. M. (1979) *Annu. Rev. Psychol.* **30**, 63–102.
22. Goel, P., Castelucci, V. F., Schacher, S. & Kandel, E. R. (1986) *Nature (London)* **322**, 419–422.
23. Squire, L. R. (1994) in *Memory Systems 1994*, eds. Schacter, D. L. & Tulving, E. (MIT Press, Cambridge, MA), pp. 203–231.
24. Squire, L. R. (1986) *Science* **232**, 1612–1619.
25. Schacter, D. L. & Tulving, E. (1994) in *Memory Systems 1994*, eds. Schacter, D. L. & Tulving, E. (MIT Press, Cambridge, MA), pp. 1–38.
26. Bendor, J. & Swistak, P. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 3596–3600.