Ni^{2+}, in silicate glasses from a tetrahedral to an octahedral environment. In addition, the band positions of octahedral Ni^{2+} move, indicating an increased electrostatic field strength around the Ni ion at high pressure. All of these structural changes are quenchable and therefore not due to simple elastic deformations.

A change of Co^{2+} from tetrahedral to octahedral coordination in a melt will increase the CFSE of this ion by about 29 kJ mol^{-1}; for Ni^{2+}, the increase in CFSE is 61 kJ mol^{-1} (ref. 13). This effect will stabilize both Co^{2+} and Ni^{2+} in silicate melts relative to crystalline silicates as well as relative to a metal phase. Crystal–silicate melt and metal–silicate melt partition coefficients will be reduced by 1 (for Co) to 2 (for Ni) orders of magnitude at 1,500 K, if a complete change from four to six-fold coordination occurs. Our data suggest that this will hold true for Co^{2+} at about 200 kbar.

This effect will not be quite as large for Ni^{2+}, because in most melt compositions only a small fraction of this species is in tetrahedral sites at low pressures. A similar effect, however, will be caused by the observed increase in field strength with pressure around the octahedrally coordinated Ni^{2+}. At high pressures, Co^{2+} and Ni^{2+} will therefore behave in a much less compatible and much less siderophile way than at low pressures. This prediction is consistent with the few available experimental data on olivine/melt partitioning at high pressures[20]. It is likely that other transition-metal ions will show similar behaviour. This means that pressure-induced coordination changes in silicate melts have to be taken into account when modelling the global chemical evolution of the Earth and the terrestrial planets Mercury, Venus and Mars. □

Received 12 November 1992; accepted 10 May 1993.

1. Irving, A. J. *Geochim. cosmochim. Acta* **42**, 743–770 (1978).
2. Hart, S. R. & Davis, K.E. *Earth planet. Sci. Lett.* **40,** 203–219 (1978).
3. Kinzler, R. J., Grove, T. L. & Recca, S. I. *Geochim. cosmochim. Acta* **54**, 1255–1265 (1990).
4. O'Neill, H. St. C. *Geochim. cosmochim. Acta* **55**, 1159–1172 (1991).
5. Ringwood, A. E. *Z. Naturforsch.* **44**a, 891–923 (1989).
6. Drake, M. J. *Z. Naturforsch.* **44**a, 883–890 (1989).
7. Colson, R.O. *Nature* **357**, 65–68 (1992).
8. Jones, J. H., Capobianco, C. J., Drake, M. J. & O'Neill, H. *Science* **257**, 1281–1285 (1992).
9. Agee, C. B. & Walker, D. *Earth planet. Sci. Lett.* **90**, 144–156 (1988).
10. O'Neill, H. St. C. *Geochim. cosmochim. Acta* **55**, 1135–1157 (1991).
11. Seifert, S., O'Neill, H. St. C & Brey, G. *Geochim. cosmochim. Acta* **52**, 603–616 (1988).
12. Burns, R. G. & Fyfe, W. S. *Science* **144**, 1001–1003 (1964).
13. Keppler, H. *Am. Miner.* **77**, 62–75 (1992).
14. Nelson, C & White, W. B. *J. Mater. Res.* **1**, 130–138 (1986).
15. Galoisy, L & Calas, G. *Am Miner.* **77**, 677–680 (1992).
16. Figgis, B. N. *Introduction to Ligand Fields* (Interscience, New York, 1966).
17. Susak, N. J. & Crerar, D. A. *Geochim. cosmochim. Acta* **49**, 555–564 (1985).
18. Greenwood, N. N. & Earnshaw, A. *Chemistry of the Elements* (Pergamon, Oxford, 1984).
19. Rossman, G. R. in *Spectroscopic Methods in Mineralogy and Geology* (ed. Hawthorne, F. C.) 207–254 (Mineralogical Society of America, Washington, 1988).
20. McFarlane, E. A. & Drake, M. J. in *Origin of the Earth* (eds Newsome, H. E. & Jones, J. H.) 135–150 (Oxford, New York, 1990).

A strategy of win–stay, lose–shift that outperforms tit-for-tat in the Prisoner's Dilemma game

Martin Nowak* & Karl Sigmund†

*Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK
†Institut für Mathematik, Universität Wien, Strudlhofgasse 4, A-1090 Vienna, Austria

THE Prisoner's Dilemma is the leading metaphor for the evolution of cooperative behaviour in populations of selfish agents, especially since the well-known computer tournaments of Axelrod[1] and their application to biological communities[2,3]. In Axelrod's simulations, the simple strategy tit-for-tat did outstandingly well and subsequently became the major paradigm for reciprocal altruism[4–12]. Here we present extended evolutionary simulations of heterogeneous ensembles of probabilistic strategies including mutation and selection, and report the unexpected success of another protagonist: Pavlov. This strategy is as simple as tit-for-tat and embodies the fundamental behavioural mechanism win–stay, lose–shift, which seems to be a widespread rule[13]. Pavlov's success is based on two important advantages over tit-for-tat: it can correct occasional mistakes and exploit unconditional cooperators. This second feature prevents Pavlov populations from being undermined by unconditional cooperators, which in turn invite defectors. Pavlov seems to be more robust than tit-for-tat, suggesting that cooperative behaviour in natural situations may often be based on win–stay, lose–shift.

Two players engaged in the Prisoner's Dilemma have to choose between cooperation (C) and defection (D). In any given round, the two players receive R points if both cooperate and only P points if both defect; but a defector exploiting a cooperator gets T points, while the cooperator receives S (with $T>R>P>S$ and $2R>T+S$). Thus in a single round it is always best to defect, but cooperation may be rewarded in an iterated (or spatial[14]) Prisoner's Dilemma.

The conspicuous success of the tit-for-tat (TFT) strategy (start with a C, and then use your co-players previous move) relies in part on the clinical neatness of a deterministic cyber-world. In natural populations, errors occur[7,12]. TFT suffers from stochastic perturbations in two ways: (1) a TFT population can be 'softened up' by random drift introducing unconditional cooperators, which allow exploiters to grow (TFT is not an evolutionarily stable strategy[15,16]); and (2) occasional mistakes between two TFT players cause long runs of mutual backbiting. (Such mistakes abound in real life: even humans are apt to vent frustrations upon innocent bystanders.)

Within the restricted world of strategies reacting only to the co-players previous move, TFT has a very important, but transitory role: in small clusters, it can invade populations of defectors, but then bows out to a related strategy, 'generous tit for tat' (GTFT), which cooperates after a co-player's C, but also with a certain probability after a D[9].

But as soon as one admits strategies which take into account the moves of both players in the previous round, evolution becomes much less transparent[17]. We first conjectured that GTFT (or variants thereof) would win the day, but are forced to admit, after extensive simulations, that the strategy Pavlov did much better in the long run. A Pavlov player cooperates if and only if both players opted for the same alternative in the previous move. The name[18] stems from the fact that this strategy embodies an almost reflex-like response to the payoff: it repeats its former move if it was rewarded by R or T points, but switches behaviour if it was punished by receiving only P or S points. This strategy, which went by the name of 'simpleton'[19], fares poorly against inveterate defectors: in every second round, it switches to cooperation. It cannot gain a foothold in a defector's world; defectors have to be invaded by other strategies, like TFT[9]. But Pavlov has two important advantages over TFT: (1) an inadvertent mistake between players using TFT causes a drawn-out battle; between two Pavlovians, it causes one round of mutual defection followed by a return to joint cooperation[26]. Thus Pavlov is fairly tolerant, like GTFT, and can correct mistakes. (2) Whereas TFT and GTFT can be invaded by drift by all-out cooperators (to the eventual profit of exploiters), Pavlov has no qualms in exploiting a sucker, once it has discovered (after an accidental mistake) that it need not fear any retaliation.

Softies cannot subvert a Pavlov population. In this sense, cooperation based on Pavlov is a safer bet than TFT. Pavlov's advantage shows best among nice strategies.

For our simulations, we consider all (stochastic) strategies with memory one (that is, recalling only the previous round). These strategies are defined by the conditional probabilities (p_1, p_2, p_3, p_4) to cooperate, given the outcome of the previous round was *R*, *S*, *T* or *P*, respectively. The game between two such strategies can be formulated as a Markov process, and its stationary distribution specifies the payoff for the infinitely iterated Prisoner's Dilemma[20]. Owing to noise, the initial move has no effect in the long run. For mathematical simplicity we retain the assumption of the infinitely iterated game, but note that the outcome is essentially unchanged if we consider sufficiently long iterated games. In our notation TFT is given by (1, 0, 1, 0) and Pavlov by (1, 0, 0, 1), but mistakes in implementing the move change 1 to $1-\varepsilon$ and 0 to ε, where ε is a small number specifying the minimal amount of noise. This is closely related to the 'trembling hand' in Selten's game theoretical notion on perfect equilibrium[15,21].

We start each simulation with the random strategy (0.5, 0.5, 0.5, 0.5). Every 100 generations (on average), we introduced a small amount of a randomly chosen mutant strategy. the frequencies of strategies spread according to the usual game dynamics[22,23], reflecting natural selection. Strategies with higher payoffs produce more offspring. Strategies whose frequency dropped below a certain threshold were discarded. Each run was observed for 10^7 generations, generating a total of about 10^5 different mutant strategies. (Note that the timescale is arbitrary, because the difference equation can be seen as an approximation of a differential equation. It is, however, very important to study the long-term dynamics and to try many mutants.) The evolutionary chronicles display an extreme diversity. Nevertheless, they allowed some clear and simple conclusions: (1) the plot for the average payoff in the population is a show-piece of punctuated equilibria (Fig. 1). Most of the time, this payoff is very close to one of the extremal values P (a regime of defection) or R (overall cooperation). The time for switching from one of the regimes to the other is usually extremely short (only a few generations). The periods of stasis frequently last for millions of generations. The later in the run, the longer they last. But the threat of a sudden collapse may never abate. (2) There is a clear tendency for cooperation (Fig. 2). After $t=10^4$ generations, only 27.5% of the runs exhibit cooperation (population payoff > 2.95); but 90% at $t=10^7$. There is also a clear tendency towards Pavlov, which dominates 10% of the runs at $t=10^4$, but

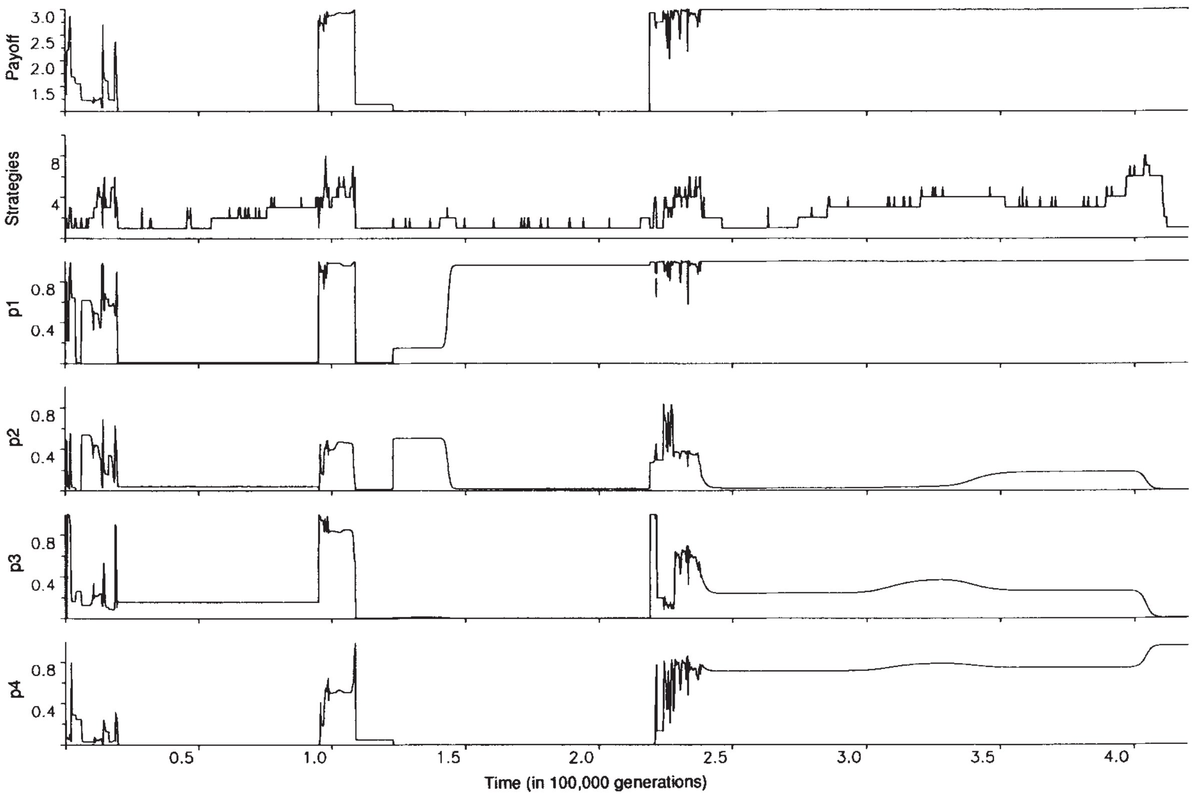

FIG. 1 An evolutionary simulation (including mutation and selection) of all strategies that consider the previous move in the iterated Prisoner's Dilemma. Such strategies are defined by four probabilities (p_1, p_2, p_3, p_4) that cooperate after having received payoff *R*, *S*, *T* or *P* in the previous round. We start with the random strategy (0.5, 0.5, 0.5, 0.5). In each generation there is a 0.01 probability that a new mutant strategy is generated at random. For technical reasons, we used the U-shaped density distribution $[\pi x(1-x)]^{-1/2}$ for sampling the p_i in the interval 0, 1. This distribution, which is familiar to geneticists, helps to explore the corners of the four-dimensional cube of stochastic strategies in a more efficient way. We assume that there is a minimal amount of noise; thus we restrict $0.001 < p_i < 0.999$. The frequencies of strategies grow according to the usual game dynamics. Strategies with frequencies below 0.001 are removed. In this example, the initial struggle for cooperation is unsuccessful. An *AllD* like strategy emerges as winner and dominates until $t \approx 92{,}000$. Then a TFT mutant invades and establishes a regime of cooperation dominated by GTFT. This population is undermined by more and more forgiving strategies (increasing p_4), which leads to a breakdown of cooperation and another period of defection, now dominated by the severe retaliator 'GRIM' (0.999, 0.001, 0.001, 0.001). Again TFT invades and catalyses the rise of cooperation. After some small adjustments in p_2, p_3, and p_4, there is a long lasting period of cooperation dominated by the Pavlov-like strategy (0.999, 0.001, 0.007, 0.946). This persists at least until $t=10^7$ (not shown). The figure shows the average population payoff, the number of strategies present at a given time, and the population averages of the probabilities p_1, p_2, p_3, and p_4. (The time is given in generations of the difference equation, but the timescale is arbitrary if the difference equation is understood as an approximation of a differential equation.) Payoff values: $R=3$, $S=0$, $T=5$, $P=1$.

82.5% at $t=10^7$. Only a few runs are eventually dominated by GTFT-like behaviour. (3) Usually, the population is monomorphic or very mildly polymorphic. We rarely find more than 10 strategies in one population. Cooperative populations (with an average population payoff close to R) are normally dominated by one or two strategies very close to Pavlov, although some long hegemonies by larger mixtures of GTFT-like strategies can also be observed. (4) Pavlov can be invaded by *AllD* if $2R<T+P$, but a prudent variant of it, (1, 0, 0, x) with $x<(R-P)/(T-R)$, is stable against *AllD*. For $2R=T+P$ (as is the case with Axelrod's payoff values), Pavlov can be invaded by *AllD*, but the stochastic, Pavlov-like strategy (0.999, 0.001, 0.001, 0.995) cannot. We observe that strategies very close to this 'almost Pavlov' win most evolutionary runs for $R=3$. Figure 3 shows the outcome of computer simulations for various different values of R. For most values of R, Pavlov clearly dominates.

How does this relate to real biology? It might well be that many cases of cooperation based on reciprocal altruism are due to Pavlov rather than to TFT. In more natural set-ups, retaliation has usually been interpreted as evidence for TFT[4-6,27]. But these experiments seem to be consistent with a Pavlov-like strategy as well. It would speak for Pavlov if animals have a tendency to exploit non-retaliators, or if they are apt to resume cooperation after bilateral defection. (In a sense, we observe Pavlov-type behaviour daily ourselves. Usually, a domestic misunderstanding causes a quarrel, after which cooperation is resumed; and the advice 'never to give a sucker an even break' is frequently adopted among members of our species.) Incidentally, in natural encounters where different rounds are not clearly separated, there is a problem in 'timing' the recovery—we may expect signals to evolve as cues.

The simple learning rule, win–stay, lose–shift, seems to be widespread and works in many other contexts. It is a particular instance of the 'law of effect', which states that animals perform more rewarded behaviours and less non-rewarded behaviours[13]. It is plausible that more sophisticated variants become established among higher animals, which stick to a move as long as a weighted payoff from the last few rounds is sufficiently rewarding (in the manner of Harley's learning rule[24]).

By and large, the iterated Prisoner's Dilemma has been seen as a story of TFT, but our results suggest that cooperation based on win–stay, lose–shift may be more robust. The success of Pavlov-like behaviour does not seem to be restricted to strategies, which only remember the last move. In other evolutionary runs, where mutations can extend the memory length, similar strategies have been found: typically, they resume cooperation after two rounds of mutual defection[25,26]. Like Pavlov, they correct for mistakes and exploit unconditional cooperators. There is, of course, no limit in the complexity of conceivable strategies. But it may be expected that the simple, natural rule of win–stay, lose–shift performs well under a variety of sophisticated conditions. Pavlov is no simpleton. □

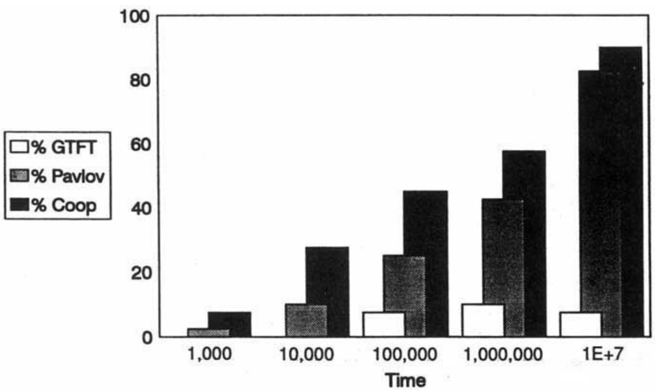

FIG. 2 There is a tendency towards cooperation and Pavlov. The figure shows the percentage of populations (at various times, arbitrary units) that are dominated by GTFT, Pavlov, or cooperative behaviour in general (population payoff average > 2.95). Note that later in the runs cooperation is always based on either Pavlov or GTFT-like behaviour. 40 simulations were performed. In each simulation $\sim 10^5$ different mutant strategies were generated at random. For each strategy the conditional probabilities to cooperate were within 0.001 and 0.999, thus the minimal amount of noise was $\varepsilon=0.001$. Pavlov-like behaviour is characterized by strategies close to (0.999, 0.001, 0.001, 0.995), and GTFT by strategies around (0.999, 0.33, 0.999, 0.33). Payoff values: $R=3$, $S=0$, $T=5$, $P=1$ (as in Axelrod's simulations).

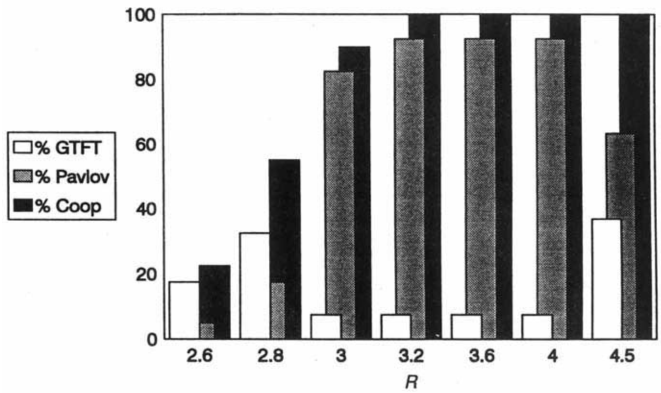

FIG. 3 Pavlov clearly outperforms TFT-like behaviour for a variety of different payoff values, R. According to the rules, R can vary between $(T+S)/2$ and T. For low values of R, GTFT seems to be slightly more successful than Pavlov, but in this case it is very difficult for cooperation to get established at all. Here the Prisoner's Dilemma is obscured by the fact that alternating C and D is almost as rewarding as cooperation. For high values of R, the populations are quite often dominated by GTFT, which for this parameter region is very close to *AllC*. For each value of R we performed 40 simulations. Pavlov-like behaviour is characterized by strategies around (0.999, 0.001, 0.001, x), where $x=0.999$ for $R>3$, $x\approx 0.995$ for $R=3$, and $x=(R-P)/(T-R)$ for $R<3$. GTFT-like behaviour is characterized by strategies around (0.999, y, 0.999, y), where $y=\min\{1-(T-R)/(R-S),\ (R-P)/(T-P)\}$. Payoff values: $S=0$, $T=5$, $P=1$ and R as indicated.

Received 1 February; accepted 15 April 1993.

1. Axelrod, R. *The Evolution of Cooperation* (Basic Books, New York, 1984).
2. Axelrod, R. & Hamilton, W. D. *Science* **211,** 1390–1396 (1981).
3. Axelrod, R. & Dion, D. *Science* **242,** 1385–1390 (1988).
4. Wilkinson, G. *Nature* **308,** 181–184 (1984).
5. Lombardo, M. P. *Science* **227,** 1363–1365 (1985).
6. Milinski, M. *Nature* **325,** 433–435 (1987).
7. May, R. M. *Nature* **327,** 15–17 (1987).
8. Dugatkin, L. A. *Behav. Ecol. Sociobiol.* **25,** 395–397 (1988).
9. Nowak, M. & Sigmund, K. *Nature* **355,** 250–253 (1992).
10. Krebs, J. R. & Davies N. B. *An Introduction to Behavioural Ecology* (Sinauer, MA, 1981).
11. Dawkins, R. *The Selfish Gene* (Oxford Univ. Press, Oxford, 1988).
12. Sigmund, K. *Games of Life* (Oxford Univ. Press, Oxford, 1993).
13. Domjan, M. & Burkhard, B. *The Principles of Learning and Behaviour* (Brooks/Cole, Monterey, 1986).
14. Nowak, M. A. & May, R. M. *Nature* **359,** 826–829 (1992).
15. Selten, R. & Hammerstein, P. *Th. Behav. Brain Sci.* **7,** 115–142 (1984).
16. Boyd, R. & Lorberbaum, J. P. *Nature* **327,** 58–59 (1987).
17. Nowak, M. & Sigmund, K. *Proc. natn. Acad. Sci. U.S.A.* **90,** 5091–5094 (1993).
18. Kraines, D. & Kraines, V. *Theory and Decision* **26,** 47–63 (1988).
19. Rapoport, A. & Chammah, A. M. *Prisoner's Dilemma* (Univ. of Michigan Press, Ann Arbor, 1965).
20. Nowak, M. & Sigmund, K. *Acta appl. Math.* **20,** 247–265 (1990).
21. Boyd, R. *J. theor. Biol.* **136,** 47–56 (1989).
22. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge Univ. Press, Cambridge, 1982).
23. Hofbauer, J. & Sigmund, K. *The Theory of Evolution and Dynamical Systems* (Cambridge Univ. Press, Cambridge, 1988).
24. Maynard Smith, J. *Th. Behav. Brain Sci.* **7,** 95–101 (1984).
25. Axelrod, R. in *Genetic Algorithms and Simulated Annealing* (ed. Davis, D.) (Pitman, London, 1987).
26. Lindgren, K. in *Artificial Life II* (eds Farmer, D. *et al.*) (Proc. Santa Fe Inst. Stud., Addison-Welsey, 1991).
27. Reboreda, J. C. & Kacelnik, A. *J. exp. Animal Behav.* **60,** 176–193 (1993).

ACKNOWLEDGEMENTS We thank W. Hamilton, A. Kacelnik, J. Krebs and R. May for discussion.