2002

Discounting and Reciprocity in an Iterated Prisoner's Dilemma

David W. Stephens
University of Minnesota - Twin Cities, dws@forager.cbs.umn.edu

Colleen M. McLinn
University of Minnesota - Twin Cities

Jeffrey R. Stevens
University of Nebraska–Lincoln, jstevens5@unl.edu

Stephens, David W.; McLinn, Colleen M.; and Stevens, Jeffrey R., "Discounting and Reciprocity in an Iterated Prisoner's Dilemma" (2002). *Jeffrey Stevens Papers & Publications*. 1.
http://digitalcommons.unl.edu/psychstevens/1

Discounting and Reciprocity in an Iterated Prisoner's Dilemma

D. W. Stephens,* C. M. McLinn, J. R. Stevens

The Iterated Prisoner's Dilemma (IPD) is a central paradigm in the study of animal cooperation. According to the IPD framework, repeated play (repetition) and reciprocity combine to maintain a cooperative equilibrium. However, experimental studies with animals suggest that cooperative behavior in IPDs is unstable, and some have suggested that strong preferences for immediate benefits (that is, temporal discounting) might explain the fragility of cooperative equilibria. We studied the effects of discounting and strategic reciprocity on cooperation in captive blue jays. Our results demonstrate an interaction between discounting and reciprocity. Blue jays show high stable levels of cooperation in treatments with reduced discounting when their opponent reciprocates, but their levels of cooperation decline in all other treatment combinations. This suggests that stable cooperation requires both reduced discounting and reciprocity, and it offers an explanation of earlier failures to find cooperation in controlled payoff games.

The Prisoner's Dilemma illustrates the economic barriers to cooperative action. In this game, the defecting (noncooperative) option is always the best choice for a single play of the game, even though both players could do better if they cooperated. Axelrod and Hamilton (1) argued that cooperation could be a game theoretical equilibrium if (i) the game was played repeatedly and (ii) the players adopted a reciprocating strategy. In their argument, repetition and reciprocity combine to make mutual cooperation a viable strategy, because although a defector will receive an immediate reward, reciprocity means that it will suffer for this choice in the long run.

Although theoreticians have exploited this paradigm with great success, it has been markedly less successful empirically (2–5). Nonhuman animals show a strong tendency to defect in experimentally created Prisoner's Dilemmas (6–9). These studies raise important questions, because we cannot usually confirm that the payoffs in naturalistic studies conform to the Prisoner's Dilemma. This uncertainty has led to controversy in some cases (10–12), and in others, it has led to questions about whether simpler explanations of observed behavior might not be more appropriate (5, 13, 14). More than 20 years after Axelrod declared the Prisoner's

Dilemma to be "the *E. coli* of social psychology" (15), there is still no single unambiguous case of stable nonhuman cooperation in a verifiable Prisoner's Dilemma.

One possible explanation for the fragility of cooperation in the Iterated Prisoner's Dilemma (IPD) is strong temporal discounting. In theory, animals should cooperate in an IPD because cooperation leads to higher payoffs in the long run, but animals may not value these long-term benefits because they strongly discount the future. Psychological studies support this idea. In these studies (16–18), experimentalists offer animals a choice between small immediate and large delayed food rewards. These experiments show very strong preferences for immediacy. Fitted discounting functions suggest that the first second of delay reduces the value of food reward by as much as 50% (19). These data, therefore, suggest that animal discounting may be much stronger than rates typically assumed by economists and other students of human behavior [e.g., 4% per year (20)].

An alternative explanation of the fragility of cooperative equilibria might hold that animals fail to cooperate, not because they discount strongly, but because they do not implement the appropriate strategy. In the IPD framework, the opponent's reciprocation means that cooperation now enhances long-
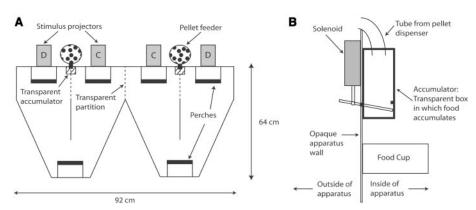
Department of Ecology, Evolution and Behavior, University of Minnesota, 1987 Upper Buford Circle, St. Paul, MN 55108, USA.

*To whom correspondence should be addressed. E-mail: dws@forager.cbs.umn.edu

Fig. 1. (**A**) Top view of apparatus. The apparatus consists of side-by-side compartments, each in the shape of a V. Each compartment is equipped with three perches. Each perch has a microswitch that reports its status to a controlling computer. Each compartment houses a single bird, one of which is designated the subject and the other is designated the stooge. The subject chooses freely, but the stooge follows an experimentally imposed strategy. At the beginning of a play, the birds wait on their respective rear perches (at the apex of the V). At a programmed time, the controlling computer switches on stimulus lights on the front panel signaling that a trial has begun. The subject may hop forward to one of the two front perches to indicate its choice. A hop on the inside perch indicates a cooperate (or "C") choice, whereas a hop on the outside perch indicates a defect (or "D") choice. The stooge only sees one stimulus light and must hop on the associated perch. The apparatus is designed with transparent partitions across the front and opaque partitions elsewhere so that the birds can see each other after they have made a choice (hopped to the front), but not before. When both birds occupy one of the front perches, the pellet dispensers deliver food into the accumulators. (**B**) Accumulator. A transparent plastic box, front and center in each compartment, received food from the pellet dispenser. The bottom of the box was a flap that could be opened by the controlling computer. Thus, during accumulated treatments, subjects could see their food gains but not consume them until the flap was opened.

term gains by increasing the likelihood of future cooperation; without such a strategy, there is no reason to cooperate, even if animals do not discount strongly.

We conducted a factorial experiment, manipulating discounting and strategy, to assess the contributions of discounting and strategic reciprocity to the fragility of the cooperative equilibrium. The subjects were 16 blue jays (*Cyanocitta cristata*) of unknown sex and mixed experimental histories. We assigned these 16 individuals to eight pairs of unrelated individuals. We designed an experimental chamber composed of side-by-side V-shaped compartments so that each member of the pair had its own compartment (Fig. 1A).
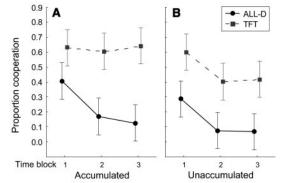
To manipulate discounting, we used payoff-accumulation treatments in which subjects had to complete a sequence of plays before obtaining the accumulated benefits from that sequence (21). Intuitively, this should emphasize the combined benefits of a sequence of interactions. To implement payoff accumulation, our apparatus dispensed food into a transparent box (Fig. 1B) so that

subjects could see their winnings but not consume them until a flap at the bottom of the box opened. In accumulated (low discounting) trials, we held food in the accumulator for four plays of the experimental game, dispensing the subject's accumulated gains at the end of the fourth play in a sequence. In unaccumulated treatments, the flap dispensed food after every play. A preliminary experiment confirmed that accumulation reduced discounting as predicted (22).

To manipulate strategies, we randomly designated one individual in a pair to be the subject and the other to be the "stooge." During data collecting trials, the subject chose freely between cooperating (C) and defecting (D) in a Prisoner's Dilemma (23). The stooge, however, played one of two experimental strategies: unilateral defection (All-D) or tit-for-tat (TFT) (cooperate initially but copy your opponent's previous move on all subsequent plays). The stooge, therefore, was forced to choose either C or D, according to its preprogrammed strategy. In broad outline, therefore, our experiment was

a conventional two-by-two factorial experiment with two levels of accumulation (accumulated and unaccumulated) and two levels of opponent strategy (All-D and TFT).

Game theoretical equilibria are stability concepts, so our design tested the effects of these four treatments on the stability of cooperation in a Prisoner's Dilemma. To achieve this, we presented each subject with a baseline matrix, which quickly led to high levels of the C response. After establishing high levels of C, we switched to a Prisoner's Dilemma payoff matrix and measured changes in the frequency of C over 1000 free-choice trials. The experiment followed a within-subjects design so that each pair experienced all four treatments in a randomly determined order (22).

Results from the four treatments are shown in Fig. 2 (22). When the opponent always defected, the frequency of cooperation declined toward zero, regardless of whether food accumulated. When the opponent reciprocated, however, we observed elevated levels of cooperation, but there was a striking difference between accumulated and unaccumulated treatments. In accumulated treatments, where in theory we reduced the effects of discounting, levels of cooperation are high and stable. In unaccumulated treatments, levels of cooperation are elevated but declining, parallel to the pattern shown in the All-D treatments. A repeated measures analysis of variance shows a significant accumulation by strategy by time interaction ($F_{2,14} = 4.31$, $P = 0.03$), supporting the idea that reduced discounting and strategic reciprocity combine to influence the stability of cooperation.
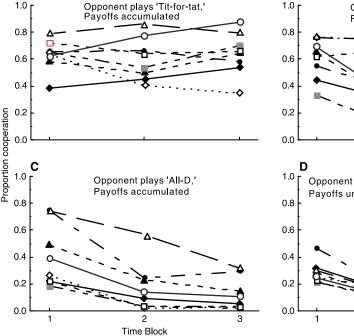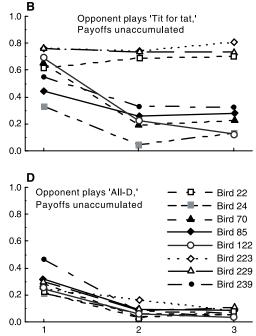
Fig. 2. Stability of cooperation in each of our four treatments. The *y* axis shows the relative frequency of the C response. The *x* axis divides trials into thirds, roughly the first 333 trials, the second 333 trials, and the third 333 trials. (**A**) Data for the accumulated (reduced discounting) treatments. (**B**) Unaccumulated (normal discounting) treatments. The dashed line shows treatments in which the opponent plays the reciprocating strategy TFT, and the solid line shows treatments in which the opponent plays All-D. Error bars indicate 95% confidence levels.

Fig. 3. Stability of cooperation showing individual variation. This is similar to Fig. 2, except that individual subjects are distinguished as shown in the legend. (**A** and **C**) Accumulated (reduced discounting) treatments. (**B** and **D**) Unaccumulated (normal discounting) treatments. In (A and B), treatments are shown in which the opponent plays the reciprocating strategy TFT, and in (C and D), treatments are shown in which the opponent plays All-D.

Table 1. Observed and theoretical strategies. The probability of the subject cooperating in the trial following the T, R, P, and S payoffs was calculated for accumulated and unaccumulated trials in the TFT treatment. A bootstrap sampling technique was used to establish confidence limits. This technique generated 1000 strategy vectors from the data. The values for TFT and Pavlov represent predicted strategy vectors for subjects implementing those strategies. Numbers in parentheses are 95% confidence limits.

Source	t	r	p	s
Data: Accumulated	0.644 (0.600, 0.683)	0.755 (0.731, 0.780)	0.394 (0.349, 0.437)	0.587 (0.545, 0.628)
Data: Unaccumulated	0.381 (0.338, 0.425)	0.686 (0.652, 0.720)	0.225 (0.202, 0.248)	0.496 (0.450, 0.542)
Theory: Tit-for-tat	1.0	1.0	0.0	0.0
Theory: Pavlov	0.0	1.0	1.0	0.0

Data in Fig. 3 are similar to those in Fig. 2, except that Fig. 3 shows separate results for each subject. Individuals were most consistent in the All-D/unaccumulated condition, where all eight subjects approached zero cooperation by the end of the treatment. In the TFT/accumulated treatment, there is more variation, but all individuals were cooperating at elevated levels at the end of the treatment. In the TFT/unaccumulated condition, we observed some intriguing bimodality, with three of eight birds showing stable levels of cooperation (as in the TFT/accumulated treatment) and the other five birds showing an erosion of cooperation that parallels the data in the All-D treatments. The elevated levels of cooperation observed in the TFT/unaccumulated treatment (Fig. 2B) are largely due to these three individuals.

Our data allow some characterizations of the subjects' strategies in response to the experimentally created reciprocity of the stooge. One can represent a variety of first-order strategies with the vector representation (t, r, p, s), where t is the probability of cooperating after obtaining payoff T [subject defected, stooge cooperated (23)], r is the probability of cooperating after obtaining payoff R (both subject and stooge cooperated), and so on (21). Using this notation, we represent TFT as (1, 1, 0, 0), All-D as (0, 0, 0, 0), and Nowak and Sigmund's (24) Pavlov as (0, 1, 1, 0). Table 1 shows calculated strategy vectors for the two treatments in which the opponent played TFT (25).

We saw three patterns. First, all four components of the strategy vector were higher in the accumulated treatment than in the unaccumulated treatment. We observed especially striking differences between the accumulated and unaccumulated treatments for the t and p components of the strategy. This means that subjects were more likely to switch from defection to cooperation when payoffs accumulated, regardless of whether their most recent defection resulted in a large (T) or mediocre (P) payoff. In contrast, the r and s components of the strategy varied relatively little, suggesting that the main effect of accu-

mulation is an increased willingness to switch from defection to cooperation. Second, the overall pattern of the observed strategy was $r > t \approx s > p$; that is, subjects were most likely to cooperate after mutual cooperation and least likely to cooperate after mutual defection, but they cooperated at roughly equivalent intermediate levels in the two "mixed" situations (t, subject defected and stooge cooperated; s, subject cooperated and stooge defected). Third, as Table 1 shows, the observed strategy disagrees strikingly with both Pavlov and TFT. For example, both Pavlov and TFT predict that $s = 0$ (do not cooperate with a player who has just "suckered" you), but our subjects were extremely forgiving, cooperating at rates near 50% after they had been suckered.

Our results show that discounting and strategy both affect the stability of cooperation; we observed high stable levels of cooperation when payoff accumulation reduced discounting and the opponent played the strongly reciprocating strategy TFT. We observed declining levels of cooperation in all other treatments. Several recent critiques of the Prisoner's Dilemma have focused on discounting (7, 9, 26, 27). Our result is in general agreement with these critiques, but it also emphasizes the complementary relation between discounting and strategy. It is, after all, the strategy that creates a pattern of future gains.

Our work suggests that the timing of benefits can be the difference between stable cooperation and cooperation that erodes to mutual defection. These results agree in some respects with Axelrod and Hamilton's influential framework (1), because they show that reciprocity combines with future value to stabilize cooperative action. Yet, our results also agree with the work of those who have challenged the general applicability of the IPD framework to real animal cooperation, because the experimental machinations required to stabilize cooperation in our study are special. Specifically, our results disagree with the field's traditional focus on simple undiscounted repetition. In contrast, they provide

solid evidence for the role of discounting and impulsivity in animal cooperation and therefore raise questions about what factors can reduce discounting enough to promote cooperation. Information about when benefits are realized in cases of natural cooperation may provide important insights into the organization of animal social behavior.

References and Notes

1. R. Axelrod, W. D. Hamilton, *Science* **211**, 1390 (1981).
2. L. A. Dugatkin, *Cooperation Among Animals: An Evolutionary Perspective* (Oxford Univ. Press, New York, 1997).
3. L. A. Dugatkin, M. Mesterton-Gibbons, A. I. Houston, *Trends Ecol. Evol.* **7**, 202 (1992).
4. A. E. Pusey, C. Packer, in *Behavioral Ecology: An Evolutionary Approach*, J. R. Krebs, N. B. Davies, Eds. (Blackwell Scientific, Oxford, UK, 1997), pp. 254–283.
5. T. H. Clutton-Brock et al., *Nature* **284**, 1640 (1999).
6. R. M. Gardner, T. L. Corbin, J. S. Beltramo, G. S. Nickell, *Psychol. Rep.* **55**, 687 (1984).
7. K. C. Clements, D. W. Stephens, *Anim. Behav.* **50**, 527 (1995).
8. M. Flood, K. Lendenmann, A. Rappoport, *Behav. Sci.* **28**, 65 (1983).
9. L. Green, P. C. Price, M. E. Hamburger, *J. Exp. Anal. Behav.* **64**, 1 (1995).
10. J.-G. J. Godin, S. A. Davis, *Proc. R. Soc. London Ser. B* **259**, 193 (1995).
11. M. Milinski, *Nature* **325**, 433 (1987).
12. D. W. Stephens, J. P. Anderson, K. E. Benson, *Anim. Behav.* **53**, 113 (1997).
13. B. Heinrich, *Behav. Ecol. Sociobiol.* **23**, 141 (1988).
14. _____, *Ravens in Winter* (Simon & Schuster, New York, 1989).
15. R. Axelrod, *J. Conflict Resolut.* **24**, 3 (1980).
16. G. Ainslie, *J. Exp. Anal. Behav.* **21**, 485 (1974).
17. H. Rachlin, L. Green, *J. Exp. Anal. Behav.* **17**, 15 (1972).
18. J. E. Mazur, in *Quantitative Analyses of Behavior: The Effect of Delay and of Intervening Events on Reinforcement Value*, M. L. Commons et al., Eds. (Erlbaum, Hillsdale, NJ, 1987), pp. 55–73.
19. J. E. Mazur, *J. Exp. Psychol. Anim. Behav. Processes* **10**, 426 (1984).
20. M. L. Weitzman, *Am. Econ. Rev.* **91**, 260 (2001).
21. D. W. Stephens, *J. Theor. Biol.* **205**, 1 (2000).
22. Materials and methods are available as supporting material on *Science* Online.
23. The subject experienced the following Prisoner's Dilemma payoff matrix: four food pellets when both subject and stooge cooperated (R, in conventional notation), zero pellets when the subject cooperated and the stooge defected (S), six pellets when the subject defected and the stooge cooperated (T), and two pellets when both individuals defected (P). The stooge always obtained two food pellets per play.
24. M. Nowak, K. Sigmund, *Nature* **364**, 56 (1993).
25. We could not calculate similar values for conditions in which the opponent played All-D, because in these cases, the subject never has the opportunity to respond to its opponent's cooperation (so t and r cannot be calculated).
26. D. W. Stephens, K. Nishimura, K. B. Toyer, *J. Theor. Biol.* **176**, 457 (1995).
27. A. Silverstein, D. Cross, J. Brown, H. Rachlin, *J. Behav. Decis. Making* **11**, 123 (1998).
28. This work was supported by the NSF and the National Institute of Mental Health. We thank E. Fernandez-Juricic, M. Hauser, B. Kerr, C. Packer, and S. Polasky, who commented on earlier drafts of this manuscript.

Materials and Methods

Here we describe materials and methods of two experiments, as well as present SOM text with brief results. The first was a preliminary experiment to explore the effects of payoff accumulation on discounting. We then provide detail on the specific methods for the Iterated Prisoner's Dilemma experiment.

General Methods

Apparatus

The apparatus used in both experiments is shown in Fig. 1A. Each V-shaped compartment was equipped with one rear perch and two front perches, all connected to microswitches so that the experimental program could detect which perch a bird occupied. Each perch was positioned immediately below a stimulus light (Med Associates ENV-123) that could display any of several colors.

The front panel of each compartment was also equipped with a food cup, an associated pellet dispenser (dispensing 20-mg pellets, Med Associates ENV-203-20), and a custom-made small transparent box in which pellets accumulated (Fig. 1B). At experimentally determined times, a flap on the bottom of this box could be opened (by energizing a solenoid), delivering the pellets into the food cup. This device allowed us to create accumulated treatments in which food accumulated, seen but unavailable, over several trials, and unaccumulated treatments in which food was dispensed immediately at the end of a trial.

Training

Before being tested in an experiment, subjects were trained to wait on the rear perch for a fixed time and then hop forward to an illuminated front perch to obtain food. This training used conventional shaping (or successive approximation) techniques and, typically, took 3 to 6 weeks to complete.

Closed Economy

We ran both experiments as a closed economy: Subjects lived in the apparatus and obtained all of their food from it. Experimental contingencies were in effect from 0700 until 1100 and from 1200 to 1600 every day. There was a 1-hour break for apparatus maintenance and animal care from 1100 to 1200 daily; otherwise, the animals stayed in the apparatus.

All stimulus lights were dark during the periods when contingencies were not in effect. To ensure that subjects obtained sufficient food to survive, we provided supplementary food (at 1600 daily) for any bird that obtained <7 g during the day. A white noise generator provided masking noise whenever the experimental contingencies were in effect.

Preliminary Experiment

The goal of the preliminary experiment was to determine whether and under what conditions accumulation and the temporal arrangement of trials influence blue jay preferences for immediacy (or discounting). This experiment followed the self-control procedure typically used in discounting studies [see (*S1*) for an example]. In self-control studies, subjects must choose between a small immediate benefit and a benefit that is larger but more delayed. Animals with strong preferences for the small immediate benefit are said to exhibit strong discounting. This experiment considered the effects of temporal clumping of trials and the accumulation of benefits across trials.

Subjects

The subjects in the preliminary experiment were six adult blue jays of unknown sex and mixed experimental histories: Band numbers were b70, b85, b108, b223, b229, and b239.

Overview of a Trial

The sequence of events within a trial was as follows. (i) The subjects waited for a fixed time (the intertrial interval). (ii) The apparatus offered a choice between small immediate and large delayed options, by switching on lights of different colors at the front of the apparatus. (iii) The subject chose one of the options by hopping forward to the associated perch. Once the choice was made, the unchosen light was switched off, and the experimentally programmed delay to

food began. (iv) When the programmed delay expired, food was dispensed, and the process began again at step 1.

Trials were organized into blocks of 32. The first 8 trials within each block were forced (or "no choice") trials in which the subjects were offered only one option (i.e., either the small immediate option or the large delayed option, but not both); the remaining 24 trials were free-choice trials. We randomly selected the light colors associated with small immediate and large delayed options for each subject, and this association was maintained throughout the experiment.

Treatments

Clumping treatments. Trials were arranged into groups of four (within the blocks of 32 discussed above). In temporally "clumped" treatments, the subject waited 345 s between clumps and then was presented with a quick succession of four trials, with a 5-s gap between each trial, so that it experienced four trials every 360 s. In unclumped trials, the subjected waited 90 s between each trial, again experiencing four trials every 360 s.

Accumulation treatments. In accumulated treatments, food was held in the accumulator—visible to the subject, but unavailable—for four trials and was delivered immediately after the fourth trial was completed. In unaccumulated treatments, the accumulator dispensed food immediately.

Delay treatments. To assess whether our manipulations influenced the blue jays' sensitivity to delay, we tested three levels of delay-to-small (one 20-mg food pellet after 5, 15, or 30 s) and two levels of delay-to-large (three 20-mg food pellets after 45 or 75 s). Each subject, therefore, experienced 24 distinct treatments (two levels of accumulation, two levels of clumping, two levels of delay-to-large, and three levels of delay-to-small). We first randomized the order in which each subject experienced the four accumulation/clumping treatments. We then randomized the order in which they received the six delay treatments within each accumulation/clumping treatment. Each delay treatment ran for 3 days, yielding ~430 free trials per treatment.

Baseline treatments. To minimize order effects, we set up the experiment so that the subjects experienced a baseline treatment before testing in each accumulation/clumping treatment. We designed the baseline treatment to be intermediate between the actual experimental treatments. The delay-to-large was 60 s, and the delay-to-small was 25 s. At the

beginning of each set of four trials, the baseline program randomly determined whether the next four trials would be (i) clumped or not and (ii) accumulated or not. In clumped trials, the subject waited 285 s before being presented with a succession of four trials, one 5 s after the other. In unclumped trials, the subject waited 75 s between each of the four trials. Subjects experienced 4 days of these baseline trials before starting each new accumulation/clumping combination.

Brief Results

Repeated measures of analysis of variance (ANOVA) of the arcsine-transformed data found a significant accumulation/clumping interaction ($F_{1,5} = 9.9683$, $P = 0.0252$), suggesting that temporal clumping enhanced the effect of accumulation on discounting. Specifically, payoff accumulation shifted individual jays' preference toward larger, more delayed options when trials were clumped in time (univariate test of significance for post hoc comparison, $F_{1,5} = 13.3131$, $P = 0.0148$).

Main Experiment

The main experiment was designed to study the interaction between strategy and discounting in an Iterated Prisoner's Dilemma. In this experiment, pairs of blue jays repeatedly chose between two alternatives representing cooperation (C) and defection (D). To manipulate strategies, we assigned one individual in each pair to act as a stooge. The stooge was trained to simply follow lights, and in doing so it could be made to follow an experimentally determined strategy [either tit-for-tat (TFT) or All-D]. We used payoff accumulation to manipulate discounting.

Subjects

The subjects were eight pairs of adult blue jays (16 birds total). In an effort to maintain similar levels of motivation, we chose birds with similar body weights for each pair. We randomly designated one member of the pair as the subject and another as the stooge. The pairs in this experiment (subjects are listed first in each pair) were b22 and b18, b24 and b84, b70 and b1, b85 and b140, b122 and b3, b223 and b106, b229 and b130, and b239 and b208. In referring to the pairs, we cite only the subject's band number.

Overview of a Trial

The sequence of events within a single trial (or play of the game) was as follows. (i) The computer switched on the rear lights at the beginning of each trial, signaling that a new trial had started. When both subjects occupied the corresponding rear perches, the rear lights were "washed out" (by switching on an additional white light) to indicate that beginning of the intertrial interval. (ii) When the intertrial interval had expired, the subject's front lights were switched on, indicating that a choice (or play) had become available. The appropriate front light (as determined by the programmed strategy) for the stooge was switched on at the same time. The trial only proceeded to this choice phase if both individuals were on their rear perches. (iii) Next, both birds hopped forward to one of the front perches, the unchosen light was extinguished, and the chosen light was washed out as described above. (iv) Once both birds occupied their front perches, food was dispensed according to experimentally determined game matrices.

Treatments

Following the results of the preliminary experiment, all trials were arranged into clumps of four to enhance the effect of accumulation on discounting. The birds waited 345 s and then played four times in quick succession with 5 s between plays. In addition, trials were arranged in blocks of 40 (10 clumps of 4), with 8 forced or no-choice trials followed by 32 free-choice trials. During the forced choice trials, the apparatus presented only one option to the subjects (either C or D), whereas the stooge continued to play its programmed strategy (TFT or All-D).

The C and D choices were defined by their positions (see Fig. 1A). In addition, we randomly assigned the color of stimulus light associated with the subject's C and D for each treatment. The stooge's lights were both the same color.

Stooge strategy treatments. As explained above, the stooge implemented a strategy that we determined (either TFT or All-D). We simply programmed the apparatus so that the appropriate choice (C or D) was the only one available to the stooge. To program TFT, for example, we programmed the apparatus to show the inside C light, if the subject cooperated on the previous trial, and to show the outside D light, if the subject defected on the previous trial. The appropriate payoff matrix (baseline or PD) determined the subject's payoffs, just as if the

stooge had chosen freely. When there were long programmed gaps between plays, such as from one day to the next or over the midday break, the TFT player began by cooperating.

Accumulation treatments. In accumulated treatments, food was held in the transparent accumulator for a series of four trials, whereas in unaccumulated treatments, the flap was opened at the end of each trial, immediately dispensing food.

Game matrices used in treatments. In describing payoff matrices, we use the conventional Prisoner's Dilemma notation. That is, a focal player obtains R (the reward for cooperation) if both players cooperate, obtains S (the sucker's payoff) if the focal player cooperates and his opponent defects, obtains T (the temptation to cheat) if the focal player defects and his opponent cooperates, and obtains P (the punishment payoff) if both players defect.

Game theoretical equilibria are stability concepts. To test the stability of cooperation in a Prisoner's Dilemma treatment, we sought to first create high levels of cooperation. To achieve this, we presented each subject with a baseline matrix ($R = 4$ pellets, $S = 2$ pellets, $T = 0$ pellets, $P = 0$ pellets) that quickly led to high levels of the C response. Subjects experienced this baseline treatment before each Prisoner's Dilemma test, for a minimum of 3 days, terminating only when the subject showed 80% C or higher for two consecutive days. All treatment variables (accumulation, stooge strategy) were in effect during this baseline phase.

During tests, the subject experienced the Prisoner's Dilemma payoff matrix $R = 4$ pellets, $S = 0$ pellets, $T = 6$ pellets, $P = 2$ pellets. The stooge obtained two pellets on all trials. Thus, the stooge received slightly less food on average than the subject did. We did this to maintain the stooge's overall hunger level and to make sure that the stooge was equally motivated to choose either the C or D alternative.

Summary of general treatment procedures. The general plan of a single treatment was as follows. (i) Stimulus colors, stooge strategy, and accumulation treatment were randomly assigned. (ii) The subject experienced the baseline matrix (with the chosen accumulation and stooge-strategy treatments in effect) until the cooperation criterion was satisfied. (iii) The subject's payoff matrix was changed to the PD matrix, and this was in effect until the subject completed 1000 free-choice trials (5 to 7 days). We repeated this cycle until each subject had completed all four accumulation/strategy combinations. In one case, a subject completed less

than 1000 free trials (never less than 800 free trials), because of an error in the program that managed the transitions from one treatment to the next.

Brief Results

A repeated measures ANOVA of arcsine-transformed choice proportions found the following effects: main effect of time period $F_{2,14} = 59.61$, $P \approx 0.0$; main effect of opponent strategy $F_{1,7} = 63.72$, $P = 0.000092$; main effect of accumulation $F_{1,7} = 4.20$, $P = 0.08$; two-way interaction of time period and opponent strategy $F_{2,14} = 25.12$, $P = 0.000023$; two-way interaction of time period and accumulation $F_{2,14} = 1.28$, $P = 0.31$; two-way interaction of opponent strategy and accumulation $F_{2,14} = 0.42$, $P = 0.54$; and three-way interaction of time period, opponent strategy, and accumulation $F_{2,14} = 4.31$, $P = 0.034$.

<ins>References and Notes</ins>

S1. D. W. Stephens, D. Anderson. *Behav. Ecol.* **12**, 330 (2001).