JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1995, 64, 1–17 NUMBER 1 (JULY)

PRISONER'S DILEMMA AND THE PIGEON: CONTROL BY IMMEDIATE CONSEQUENCES

LEONARD GREEN, PAUL C. PRICE, AND MERLE E. HAMBURGER

WASHINGTON UNIVERSITY, UNIVERSITY OF MICHIGAN, AND STATE UNIVERSITY OF NEW YORK AT ALBANY

In three experiments pigeons played (i.e., chose between two colored keys) iterated prisoner's dilemma and other 2 × 2 games (2 participants and 2 options) against response strategies programmed on a computer. Under the prisoner's dilemma pay-off matrix, the birds generally defected (i.e., pecked the color associated with not cooperating) against both a random response (.5 probability of either alternative) and a tit-for-tat strategy (on trial n the computer "chooses" the alternative that is the same as the one chosen by the subject on trial $n - 1$) played by the computer. They consistently defected in the tit-for-tat condition despite the fact that as a consequence they earned about one third of the food that they could have if they had cooperated (i.e., pecked the "cooperate " color) on all the trials. Manipulation of the values of the food pay-offs demonstrated that the defection and consequent loss of food under the tit-for-tat condition were not due to a lack of sensitivity to differences in pay-off values, nor to strict avoidance of a null pay-off (no food on a trial), nor to insensitivity to the local (current trial) reward contingencies. Rather, the birds markedly discounted future outcomes and thus made their response choices based on immediate outcomes available on the present trial rather than on long-term delayed outcomes over many trials. That is, the birds were impulsive, choosing smaller but more immediate rewards, and did not demonstrate self-control. Implications for the study of cooperation and competition in both humans and nonhumans are discussed.

Key words: prisoner's dilemma, 2 × 2 games, tit for tat, self-control, cooperation, key peck, pigeons

---

Dubbed the *E. coli* of social psychology by Axelrod (1980a), the prisoner's dilemma game has become a popular research tool and model of social interactions. Conflict situations from the interpersonal (e.g., Baefsky & Berger, 1974; Baum & Gatchel, 1981) to the international (Snyder & Diesing, 1977) have been modeled, with varying degrees of success, as prisoner's dilemma games.

The prisoner's dilemma game (actually one class of 2 × 2 games) involves two players who must each choose between two response alternatives, generally called cooperation and defection, without knowing what the other's choice will be. Depending on the combination of choices made by the two players, each receives one of four possible pay-offs—often denoted S, P, R, and T—that satisfy the inequalities $S < P < R < T$, and $R > (T + S)/2$. The pay-off structure of the prisoner's dilemma game is represented by the matrix in Figure 1. The rows of the matrix represent Player 1's possible choices, and the columns represent Player 2's possible choices. The cell formed by the intersection of a row and column represents the outcome of that combination of choices and shows the resulting pay-offs (assumed to be utilities), the first of which is that earned by Player 1, and the second of which is that earned by Player 2. When both players cooperate, each receives a moderate pay-off, R, the "reward" for mutual cooperation. When both players defect, each receives a smaller pay-off, P, the "punishment" for mutual defection. When one player cooperates and the other defects, the defecting player earns the highest possible pay-off, T—the "temptation" to defect—and the cooperating player earns the lowest possible pay-off, S, commonly (and inappropriately) known as the "sucker's" pay-off. Each player in the prisoner's dilemma, regardless of what the other chooses, earns a greater pay-off by defecting than by cooperating. (With reference to Figure 1, defection as compared with cooperation would yield five rather than three should the opponent cooperate, and one rather than zero should the opponent defect.) In game-theory terms, defection is said to dominate cooperation. The dilemma arises because if both players defect, each re-

---

Our thanks to members of the Psychonomy Cabal; in particular, David Hsu for assistance with Experiment 3, Debra Freed for her suggestions and extensive work, and Joel Myerson for superb comments on drafts of the paper.

Address correspondence to Leonard Green, Department of Psychology, Washington University, Campus Box 1125, St. Louis, Missouri 63130.

		Player 2: cooperate	Player 2: defect
Player 1	cooperate	3, 3 (R, R)	0, 5 (S, T)
	defect	5, 0 (T, S)	1, 1 (P, P)

Fig. 1. Pay-off matrix for the prisoner's dilemma game. In any given cell, the left number is the outcome for Player 1, and the right number is the outcome for Player 2. Outcomes are in terms of positive utility.

ceives a pay-off of lower utility than if they both cooperated (i.e., a pay-off of one rather than three).

It is important to distinguish between a one-shot prisoner's dilemma (in which the players meet only once) and iterated games (in which the players meet in a series of prisoner's dilemmas). Because defection dominates cooperation, some consider it to be the rational response in one-shot prisoner's dilemmas. In iterated games, however, the rational response—if rational implies optimal—depends on the strategy of the other player. If Player 2's strategy does not depend upon Player 1's choices (is noncontingent), then it is in Player 1's long-term best interest to defect on every trial. If Player 2 uses a strategy of permanent retaliation, cooperating on every trial until Player 1 defects and then defecting on every trial thereafter, then Player 1 maximizes his or her long-term pay-off by cooperating on every trial (with the exception of the last trial if it is known which is to be the last). For this reason, the iterated prisoner's dilemma poses a more interesting problem than the one-shot version, both for players and for researchers. The present study concerns this form of the prisoner's dilemma.

Recently, the prisoner's dilemma and other game-theory models have been used to account for cooperative and competitive behavior in nonhuman animals. In general, this line of theorizing has been concerned with the question of how cooperative behavior could have evolved among populations of organisms that were initially noncooperative, without invoking the concepts of group selection or kin selection (Maynard Smith, 1982). One answer to this question has been that there is a particular pattern of play in the prisoner's dilemma, commonly known as a tit-for-tat strategy, that leads to mutual cooperation and according to Axelrod and Hamilton (1981) is evolutionarily stable. That is to say, the tit-for-tat strategy can gain a foothold and survive in initially noncooperative environments, and once established no individual playing a different strategy can do as well.[1] The tit-for-tat strategy dictates that an organ-

[1] However, see Selten and Hammerstein's (1984) explanation of why it is not strictly true that TFT is an evolutionarily stable strategy, Boyd and Lorberbaum's (1987) argument for why no pure strategy can be evolutionarily stable in the prisoner's dilemma game, and Nowak and Sigmund's (1993) demonstration of how a win-stay, lose-shift strategy (called "Pavlov") outperforms TFT.

ism cooperate on the trial after an opponent has cooperated and defect on the trial after an opponent has defected. Organisms that play a tit-for-tat strategy receive higher pay-offs overall in terms of their inclusive fitness than those that do not. Over time, then, tit for tat should, in theory, come to dominate the behavior of the organisms in question. Lombardo (1985) has used a tit-for-tat analysis as a model of the parenting behavior of tree swallows in the wild, and Milinski (1987) and Dugatkin (1988) have argued for the use of a tit-for-tat analysis in their work with three-spined sticklebacks and guppies, respectively. However, such analyses assume, but in no way demonstrate, the pay-off matrix necessary for tit for tat, and are open to alternative explanations (e.g., Lazarus & Metcalfe, 1990; Masters & Waite, 1990).

In a related line of research, nonhuman animals have been placed in generalized prisoner's dilemmas as a method of studying the patterns of behavior that evolve not over thousands of generations but over several trials in the lifetime of one player. This approach was first advocated by Rapoport, Guyer, and Gordon (1976) because it would have "the advantage of permitting iterated games of practically any length and also [reflect] behavior which is probably motivated by nothing other than the payoffs" (p. 129). In this vein, Gardner, Corbin, Beltramo, and Nickell (1984) released pairs of rats into two T mazes. By turning in one direction at the end of the maze, each rat could "cooperate"; by turning in the other direction, each could "defect." The T mazes had a common Plexiglas wall that allowed each rat to see the other and to see the other's choice. The rats received their pay-offs (various amounts of food) in goal boxes at the end of the maze. In both this condition and one in which the Plexiglas wall was covered with cardboard so the rats could not see each other, mutual defection was by far the most common joint response. Flood, Lendenmann, and Rapoport (1983) used two adjoining operant conditioning chambers to pit rats against one another in the prisoner's dilemma and three other 2 × 2 games. The adjoining wall of the chambers was made of coarse wire mesh that allowed the rats to see, hear, and smell one another. Each rat could cooperate or defect by pressing one of two levers in its chamber, and the pay-offs came in the form of different delays to a food reinforcer (shorter delays were assumed to have greater utility). In the prisoner's dilemma condition, near fixations on defection and one complete fixation were observed in 8 of the 12 subjects, with only 1 responding cooperatively. Both the Gardner et al. and Flood et al. studies are difficult to interpret, however, because the strategy of the subject's opponent was not controlled. Because the strategy of one player determines the optimal choice of the other, it is not possible to draw any conclusions from these studies about whether individual rats respond optimally in a generalized prisoner's dilemma.

In the present study, therefore, we attempted to see how an animal behaves in an iterated prisoner's dilemma game, while we manipulated what is arguably its more interesting variable—the strategy of the other player. Because the optimal choice when playing an iterated prisoner's dilemma game depends on both the pay-off matrix and the strategy being used by the other player, controlling and varying these outcomes and strategies is the way to determine whether an animal is behaving so as to maximize its cumulative pay-off. Against an opponent who defects on every trial or who cooperates on every trial, for example, pure defection would be optimal. In fact, pure defection is optimal against an opponent playing any noncontingent strategy, that is, a strategy that is not affected by the prior choices of the other player. Against certain contingent strategies, however, pure cooperation or a mix of cooperation and defection may be optimal.

As used in describing the present experiments, a strategy can be likened to a schedule of reinforcement. Both are rules for the delivery of consequences. Delivery of those consequences is determined by the past and present behavior of the organism and the strategy (schedule) of the opponent. Pay-off matrix values are analogous to reinforcer magnitudes. In social situations, each individual's behavior represents a strategy in that one individual's behavior serves as a schedule of reinforcement for the other's behavior. A question of importance, and one that is the focus of the present experiments, is what aspects of an opponent's strategy control the behavior of an organism.

The pigeons in the present study played iterated prisoner's dilemma and other 2 × 2 games against strategies programmed on a computer. One was a tit-for-tat (TFT) strategy, which dictates cooperation on the first trial and on subsequent trials responds as its opponent did on the previous trial. TFT is a strategy based on the notion of reciprocity, and the optimal choice in such an iterated prisoner's dilemma game would be to cooperate always. In computer tournaments, the TFT strategy was found to be extremely effective, earning a total average pay-off greater than that earned by any of its opponents (Axelrod, 1980a, 1980b).

In addition to TFT, a second strategy investigated in the present study was a random response (RND) strategy, which had an equal probability of cooperating or defecting on any trial, independent of the pigeon's choice. The rational choice in this case is to defect always because, in the long term, defecting yields a higher pay-off than cooperating does, regardless of what RND chooses. Thus, if the pigeons are sensitive to the overall pay-off structure used, then they should cooperate when playing against TFT and defect when playing against RND. On the other hand, if the pigeons' choices are controlled by the immediately occurring trial outcomes, then they should defect against both strategies. Defection would earn the greatest immediate reward on each individual trial as well as the greatest overall total reward when played against RND. However, when played against TFT, defection would earn the immediately larger reward on each individual trial but a much lower total reward over the series of trials. How would the pigeon solve the dilemma?

EXPERIMENT 1

METHOD

Subjects

Two female White Carneau pigeons (22W and 25), with previous experience pecking response keys for mixed grains, were maintained at 80% of their caged free-feeding weights via supplementary feeding following experimental sessions. Water and grit were continuously available in their home cages.

Apparatus

The experimental space was a Coulbourn Instruments pigeon chamber measuring 25 cm across the front panel, 27.5 cm from front to back, and 30 cm in height. The chamber was housed in a sound- and light-attenuating enclosure located in a room separate from that which contained the Apple IIc® computer that controlled and recorded experimental events.

The front panel of the chamber contained two response keys, a triple cue lamp, a houselight, and an opening that provided access to a pellet feeder. The response keys were 2.5 cm in diameter and were aligned vertically on the left third of the panel. They were 5 cm apart center to center, with the center of the lower key 19 cm from the floor. Each key could be transilluminated with red and green light and required a force of at least 0.2 N to operate. The center third of the panel contained both the triple cue-lamp module and the feeder opening. The three cuelights were 8 cm from the ceiling and were aligned horizontally, 2 cm from center to center. The cuelights were type 1829 bulbs covered with a yellow, blue, or white plastic cap. Only the blue cuelight and yellow cuelight were used. The rectangular feeder opening (5.5 cm high by 5 cm across) was located 10 cm below the cuelights and 7 cm from the floor. Food deliveries consisted of differing numbers of 45-mg Noyes food pellets and were accompanied by illumination of a type 1829 bulb located within the opening of the feeder. All other sources of illumination were extinguished during food deliveries. The houselight was centered on the right third of the front panel, 2 cm from the ceiling, with the light deflected upward.

Procedure

The pigeons were first trained to peck the red and green keys by an autoshaping procedure and were then placed directly on the first experimental condition. Experimental sessions were conducted daily and consisted of 55 outcome trials, the first five of which were warm-ups and were not included in the data collection. Each trial was 25 s in length and began with the illumination of the houselight for a minimum of 2 s, after which the red and green response keys were illuminated for a maximum of 10 s or until the pigeon pecked one of the keys. Whether the top key was red and the bottom key green, or vice versa, was determined randomly on each trial. When the pigeon pecked one of the two keys, that key remained illuminated for another 6 s while the other key was extinguished. During the final 3 s, either the blue or yellow cuelight was also illuminated, after which the response key, cuelight, and houselight were extinguished for 7 s during which the feeder was illuminated and the appropriate number of food pellets was delivered, one every 0.5 s. If the bird did not peck one of the response keys within 10 s of their illumination, then all lights in the chamber were extinguished for the remainder of the trial. To maintain a constant 25-s trial duration, the latency to peck was subtracted from the 10 s maximum time permitted, and the remaining time was then added to the beginning of the next trial.

The computer was programmed to play TFT and RND strategies. In the TFT condition, the computer responded to the pigeon by making the same choice as that made by the pigeon on the preceding trial. Thus, if the pigeon had "cooperated" on the fourth trial, the computer would "cooperate" on the fifth trial. The only exception to this strategy was the first trial of each session, on which the computer always cooperated. In the RND condition, the computer randomly cooperated on half the trials and defected on half the trials, independent of the pigeon's choices.

The pigeon cooperated by pecking the red response key and defected by pecking the green response key. The blue cuelight signaled cooperation on the part of the computer and the yellow cuelight signaled defection. The pay-off matrix for TFT and RND is shown in Table 1 (under prisoner's dilemma). Cooperation by the pigeon (selection of the red key) and cooperation by the computer (illumination of the blue cuelight) resulted in a three-pellet pay-off. If the pigeon defected (selected the green key) and the computer cooperated (illumination of the blue cuelight), the pigeon earned five pellets. If the pigeon cooperated (selected the red key) and the computer defected (illumination of the yellow cuelight), the pigeon earned zero pellets. If both the pigeon and the computer defected (selection of the green key by the pigeon, accompanied by illumination of the yellow cuelight), the pigeon earned one food pellet.

It is important to note that one cannot average across the row totals of Table 1 to predict the pay-offs to be received by the pigeon, except when the pigeon is playing against the RND strategy. That is, the pigeon's obtained pay-off for a cooperation is *not* equal to the average of the two possible pay-offs, namely the pay-off for a cooperation given a computer cooperation and the pay-off for a cooperation given a computer defection. Likewise, the pigeon's obtained pay-off for a defection is, in general, not equal to the average of the two possible pay-offs given the two possible computer responses. This is because, with the exception of the RND strategy, the relative frequencies with which the pigeon receives the two possible pay-offs for a given response depends on its pattern of responding.

Each bird was studied against both strategies, with Bird 25 playing against TFT and Bird 22W playing against RND to begin the experiment. Furthermore, the meanings of the response key and cuelight colors were reversed several times. In Condition 1, the red

Table 1

Pay-off matrices for the different 2 × 2 games studied in the present experiments. Only the pay-offs (in number of pellets) earned by the pigeon are presented. C represents cooperate; D represents defect.

		Prisoner's dilemma (TFT/RND)		TFT-dom		TFT + 1		Chicken		TFT × 2	
		Computer		Computer		Computer		Computer		Computer	
		C	D	C	D	C	D	C	D	C	D
Pigeon	C	3	0	3	2	4	1	3	1	6	0
	D	5	1	0	1	6	2	5	0	10	2

keylight and the blue cuelight represented cooperation by the pigeon and the computer, respectively, and the green keylight and the yellow cuelight represented defection by the pigeon and the computer, respectively. In Condition 2, the meanings of these colors were reversed so that green and yellow represented cooperation and red and blue represented defection. In Condition 3, the colors were returned to their original meanings. In Condition 4, Bird 25 was transferred to the RND condition and Bird 22W was transferred to the TFT condition. The meanings of the keylight and cuelight colors were again reversed in Condition 5, and were changed back again to their original meanings in Condition 6.

Both birds were then studied against a TFT-dom strategy, in which the computer played a tit-for-tat strategy, but with the pay-off structure of the game altered (as shown in Table 1) so that cooperation was now the dominant response. That is, no matter what choice the computer made, the pigeon would maximize its current pay-off by cooperating.[2] If the pigeon cooperated and the computer cooperated in this condition, the bird earned three food pellets as in the previous condition. If the pigeon defected and the computer cooperated, the bird earned zero pellets. If the pigeon cooperated but the computer defected, the bird earned two pellets. If both the pigeon and the computer defected, then, as before, the bird earned one pellet. Following this condition (Condition 7), the meanings of the keylight and cuelight colors were reversed (Condition 8), followed by a return to their original meanings (Condition 9).

Conditions 10 and 11 were a return to the original pay-off matrix with both pigeons studied under the TFT contingencies. Condition 11 differed from Condition 10 only in that the meanings of the keylight and cuelight colors were reversed once more.

Table 2 presents the order of the experimental conditions. Conditions changed when both birds met the following criteria: They completed a minimum of 15 sessions, and the last five sessions were judged by two observers to be stable with no apparent trend in either bird's choice responding. Conditions were studied for an average of 30 sessions each.

[2] Note that the game played under TFT-dom is technically no longer a prisoner's dilemma game because its pay-off values do not satisfy the inequalities previously stipulated. We continue to use the terms *cooperate* and *defect* in discussing TFT-dom, however, for the sake of simplicity.

Table 2

Order of experimental conditions and the median number of cooperations and pellets earned from the last five sessions of each condition for each pigeon in Experiment 1.

Bird	Condition[a]	Number of cooperations[b]	Number of pellets earned	Maximum	Minimum
25	1. TFT	0	50	150	50
	2. TFT-rev	6	68	150	50
	3. TFT	6	68	150	50
	4. RND	4	140	150	75
	5. RND-rev	8	136	150	75
	6. RND	9	133	150	75
	7. TFT-dom	49	146	150	25
	8. TFT-dom-rev	47	137	150	25
	9. TFT-dom	50	150	150	25
	10. TFT	2	55	150	50
	11. TFT-rev	4	62	150	50
22W	1. RND	1	149	150	75
	2. RND-rev	1	145	150	75
	3. RND	7	139	150	75
	4. TFT	1	53	150	50
	5. TFT-rev	50	150	150	50
	6. TFT	0	50	150	50
	7. TFT-dom	50	150	150	25
	8. TFT-dom-rev	50	150	150	25
	9. TFT-dom	50	150	150	25
	10. TFT	0	50	150	50
	11. TFT-rev	1	53	150	50

[a] TFT represents the tit-for-tat strategy; RND represents the random strategy; rev represents a reversal in the meaning of the colors; dom represents a change in the payoff matrix of the tit-for-tat strategy to one in which the dominant response is to cooperate.

[b] The number of cooperations could vary between 0 and 50.

Results

The median numbers of cooperations and pellets earned from the last five sessions on each condition for each pigeon are shown in Table 2. Also presented are the maximum and minimum number of pellets possible in each condition. Against the TFT and RND conditions (Conditions 1 through 6), the birds defected on most of the trials. Bird 25 defected under the TFT and RND conditions and continued to defect when the meanings of the colors were reversed and then returned to their original meanings. Bird 22W defected on almost all the trials under all three of the RND conditions and continued to defect when transferred to the TFT condition. However, when the colors were reversed, Bird 22W demonstrated an apparent color bias (continuing to peck the red key), thus cooperating; when the colors were returned to their original meanings, the bird defected on all the trials.

In Conditions 7 through 9, the pay-off matrix was changed so that now the dominant response was to cooperate. Both birds under the TFT-dom condition came to cooperate on almost all trials, even when the meanings of the colors were reversed (Condition 8) and when they were returned to their original meanings (Condition 9). When then placed on the original pay-off matrix (Condition 10), both birds under the TFT condition returned to their previous behavior of defecting on almost all trials and continued to defect when the meanings of the colors were reversed once more (Condition 11). The change in behavior from cooperation under TFT-dom to defection under TFT also produced a decrease in number of pellets earned from the maximum of 150 to close to the minimum of 50.

Discussion

Under the TFT and RND conditions, the birds' choices approached pure defection. In only one of the five TFT conditions, and only for 1 bird (Condition 5, TFT-rev, Bird 22W), did cooperation predominate. Upon reversal of the meaning of the colors and in all subsequent TFT conditions, however, this bird defected, supporting the supposition that this one instance of cooperation was a temporary color bias. Thus, when playing against the RND condition, the pigeons' choices were nearly optimal—almost all defections—earning close to the maximum number of food pellets. When playing against the TFT condition, on the other hand, the pigeons behaved far from optimally; they overwhelmingly defected when it would have been in their long-term interest to cooperate. Consequently, they earned close to the minimum number of food pellets.

Why did the pigeons not cooperate against TFT? To expect the pigeons to cooperate against TFT, it must first be assumed that the psychological values (utilities) associated with the pay-offs, in addition to their numerical values, satisfy the inequalities of the prisoner's dilemma game. Of course, this may not have been the case. It may have been, for example, that the pigeons were not sensitive to the differences among the pay-offs. If S, P, R, and T were approximately psychologically equivalent for the pigeons, then all possible choices would have been equivalent too, and there would have been no reason to expect pure cooperation. However, there also would be no reason to expect the pure defection that was observed, thus arguing against psychological equivalence of the pay-offs as a likely explanation.

Moreover, the illumination of the cuelight that signified the computer's response had a clear effect on the birds' behavior. For example, in the first condition, the blue cuelight was correlated with a larger number of food pellets than was the yellow cuelight (three or five pellets compared to zero or one pellet). In this, as well as in the relevant conditions of the later experiments, the pigeons were observed to behave differentially to these cuelights: sustained and vigorous pecking at or grasping within their beaks of the blue light, but little, if any, pecking at the yellow light. This consistently observed finding further suggests that the birds were sensitive to the difference in pay-offs.

In addition, results from the TFT-dom condition demonstrate that the birds were sensitive to the differences in pay-off values. Because of the pay-off matrix in TFT-dom, the pigeon would maximize its current payoff by cooperating. Clearly, then, if the birds are sensitive to pay-off differences in the range of zero to five food pellets, they should engage in pure cooperation against TFT-dom, which, in fact, they did.

To expect the pigeons to cooperate when playing against TFT, it must be assumed that they integrate over several trials of the iterated prisoner's dilemma game. That is to say, not only must they remember the outcomes of previous trials, but they also must be sufficiently sensitive to potential outcomes more than one trial in advance. It could be, however, that the birds, although in fact playing what we call an iterated prisoner's dilemma game, were responding to each trial as if it were a discrete event. Thus, defection would indeed be the optimal response on each trial. This is so because on a single trial, defection (as opposed to cooperation) would earn five rather than three food pellets were the computer to cooperate, and would earn one rather than zero pellets were the computer to defect. Notice also that after a pigeon defected on TFT and earned either one or five pellets on trial *t*, there was a disincentive for the bird to cooperate on trial $t + 1$ because it would then receive the sucker's pay-off—no food pellets. The fact that three pellets would be earned on trial $t + 2$, and on all subsequent trials if the bird continued to cooperate, may not have controlled the birds' choices, especially given their steep discounting of delayed outcomes (Kagel & Green, 1987). Consequently, it is not unreasonable to assume that future trials, because of their temporal distance from the current trial, were not taken into account; immediate rewards determined choice.

It is obviously quite important to establish the birds' sensitivity to differences in pay-off values in the range of zero to five pellets, and the fact that the birds overwhelmingly cooperated under the three TFT-dom conditions seems to be evidence for this sensitivity. Consider the following possibility, however: The birds were only sensitive to whether or not they received a positive pay-off. That is, they could distinguish S, a pay-off of zero pellets, from P, R, or T, but they could not distinguish P, R, or T from one another. Thus, although a normative prediction of the birds' behavior against TFT would fail, the avoidance of a zero pay-off would explain the behavior against both the TFT and the TFT-dom conditions. Against both, the birds preferred whichever alternative never resulted in a pay-off of zero. This possibility, along with further indirect tests of temporal discounting, were tested in Experiment 2.

EXPERIMENT 2

The second experiment examined whether the results of Experiment 1 were due to the pigeons' avoidance of the zero pay-off. Such behavior on the part of the pigeon would lead to defection when playing against the RND and the TFT computer strategies but would lead to cooperation against the TFT-dom strategy. To this end, the pay-off matrix of TFT was changed by incrementing all the numbers of pay-offs by one food pellet. If the birds are sensitive to the different numbers of pellets delivered, continued defection would argue against their having responded so as to avoid the zero pay-off. In addition to this change, the dominant strategy, as well as the original TFT strategy, were again tested.

To strengthen further the conclusion that the pigeons were more likely to play the prisoner's dilemma game (as constructed here) only one trial in advance, and that their behavior was sensitive to the local contingencies, a new condition, "chicken," was studied. The idea behind chicken is that mutual defection leads to the worst outcome for both players. The situation is likened to that of two teen-age drivers heading toward each other in their cars; the one who swerves aside first is "chicken." Yet if neither swerves, both lose. In the game of chicken, constructed in the present experiment such that the computer responds as the pigeon did on the preceding trial, mutual cooperation yields a pay-off of three pellets (as was true under TFT), but mutual defection yields a pay-off of zero pellets. If the pigeon cooperates and the computer defects, the pigeon earns one pellet; if the pigeon defects and the computer cooperates, the pigeon earns five pellets. Under chicken, pure defection produces the lowest possible pay-off; defection does not completely dominate cooperation on each play of the game as it does under TFT. The optimal behavior against chicken is either pure cooperation, thus earning 150 pellets per game, or alternating between cooperation and defection, thus earning one and five pellets, respectively, totaling 150 per game. If the pigeons are only playing one trial in advance, as the previous results suggest, then they might be expected to intersperse defection with cooperation rather than always to cooperate. Under chicken there is still the incentive to defect after a cooperation (thereby earning five pellets by defecting rather than three by cooperating). But unlike TFT, there is also an incentive to switch back to cooperation following a defection because such a response produces one pellet, whereas continued defection produces none.

Finally, each entry in the pay-off matrix was doubled under TFT to test whether defection would continue when substantially greater pay-offs are provided. There is some evidence (although it is inconclusive) in the literature on prisoner's dilemma with humans that large pay-offs engender greater levels of cooperation (see, e.g., Insko et al., 1993).

METHOD

Subjects and Apparatus

The same 2 pigeons (22W and 25) and the same apparatus from Experiment 1 were used.

Procedure

All details of procedure were the same as those in Experiment 1. The first experimental condition studied was the final TFT strategy from Experiment 1, but with the response-key and cuelight colors returned to their original meanings. The second condition was also a TFT condition, but with the pay-off matrix changed such that each outcome was incremented by one pellet from that under the original TFT strategy (referred to as TFT + 1) (see Table 1). The meanings of the keylights and cuelights were reversed for Condition 3 (TFT + 1-rev), followed by a return to their original meanings in Condition 4 (TFT + 1).

In Condition 5, the pay-off structure was changed to that of chicken (see Table 1). As under TFT, if the pigeon and the computer cooperate, the bird earns three food pellets; if the pigeon defects and the computer cooperates, the bird earns five pellets. However, if the pigeon cooperates and the computer defects, the bird earns one pellet, whereas if both pigeon and computer defect, the bird earns none.

In Condition 6, the birds were studied against the TFT-dom condition, followed by the original TFT strategy in Condition 7. In Condition 8, the pay-off matrix under TFT was doubled (TFT × 2), with the meanings of the keylights and cuelights reversed for Condition 9 (TFT × 2-rev). Table 3 presents the order of the experimental conditions. Each condition was studied for an average of 36 sessions.

Table 3

Order of experimental conditions and the median number of cooperations and pellets earned from the last five sessions of each condition for each pigeon in Experiment 2.

Bird	Condition[a]	Number of cooperations[b]	Number of pellets earned	Maximum	Minimum
25	1. TFT	3	63	150	50
	2. TFT + 1	2	102	200	100
	3. TFT + 1-rev	0	100	200	100
	4. TFT + 1	5	114	200	100
	5. Chicken	25	126	150	0
	6. TFT-dom	49	146	150	25
	7. TFT	2	56	150	50
	8. TFT × 2	0	100	300	100
	9. TFT × 2-rev	1	106	300	100
22W	1. TFT	2	56	150	50
	2. TFT + 1	3	109	200	100
	3. TFT + 1-rev	3	109	200	100
	4. TFT + 1	1	103	200	100
	5. Chicken	25	117	150	0
	6. TFT-dom	49	146	150	25
	7. TFT	2	56	150	50
	8. TFT × 2	1	106	300	100
	9. TFT × 2-rev	3	118	300	100

[a] TFT represents the tit-for-tat strategy; TFT + 1 represents the tit-for-tat strategy with all outcomes increased by one pellet; dom represents a change in the payoff matrix of the tit-for-tat strategy to one in which the dominant response is to cooperate; TFT × 2 represents the tit-for-tat strategy with the number of pellets per outcome doubled; rev represents a reversal in the meaning of the colors.

[b] The number of cooperations could vary between 0 and 50.

RESULTS AND DISCUSSION

The median numbers of cooperations and pellets earned from the last five sessions on each condition for each pigeon are presented in Table 3. Also shown are the minimum and maximum number of pellets possible under each of the conditions. As before, both birds defected on almost all of the trials when playing against the TFT strategy, earning close to the minimum number of pellets. When the number of pellets for each outcome was then incremented by one, the birds continued to defect on almost all the trials, continuing to earn close to the minimum. Defection continued when the meanings of the colors were reversed in Condition 3 and when they were then returned to their original meanings (Condition 4). Clearly, avoidance of a zero pay-off cannot explain the results of these experiments.

When the computer strategy was changed to Chicken, Birds 25 and 22W cooperated on half the trials, earning 84% and 78% of the maximum number of pellets, respectively. When then placed on the TFT-dom strategy (Condition 6), each came to cooperate on just about every trial, earning near the maximum. Upon return to the original TFT strategy, the birds markedly reduced their cooperations, defecting on all but two trials each. When the number of pellets per outcome was doubled (Condition 8), the birds continued to defect on almost all trials. Defection continued under Condition 9 when the meanings of the colors were reversed.

The pattern of these results suggests that (a) the birds were sensitive to the pay-off values used, (b) the birds altered their choices according to the strategy of the computer opponent, (c) the TFT results were not due to only one set of pay-off values having been studied, and (d) the pigeons did not respond to the overall pay-off structure, but instead were overly influenced by the pay-off of the immediately occurring trial. The birds' marked discounting of delayed outcomes is even more apparent when one considers that upon a reversal in the meaning of the colors under TFT, the birds initially continued to peck the same key color as they had during the final days of the previous condition, consequently tripling their number of pellets earned. Yet they subsequently changed to pecking the other color and thus decreased their pellets earned from the maximum to close to the minimum. This is illustrated in Figure 2, in which the final 5 days on the TFT condition are presented followed by the 31 days on the TFT-rev condition (Conditions 10 and 11 from Experiment 1). Plotted are the number of choices of the red key and the number of food pellets earned. During TFT, red meant cooperate; during the reversal, red meant defect. In the TFT condition, both birds were defecting (pecking green) and earning near the minimum number of pellets. When the meanings of the colors were then reversed, each bird initially continued to peck the green key, nearly tripling its earnings. Yet, after five or 16 sessions for Birds 25 and 22W, respectively, choice shifted to the red (defect) key and the number of pellets earned dropped from the maximum to near minimum (from 150 food pellets in the first session of TFT-rev to 56 and 50 food pellets in the final session). A similar analysis holds for the change from TFT-dom to TFT.

In order to make comparisons across the various conditions from Experiments 1 and 2 in terms of how the choices of the pigeons influenced their earnings, taking into account that there were different minimum and maximum numbers of pellets that could be earned in the various conditions, the percentage of the difference between the minimum and the maximum number of pellets possible that was actually earned was calculated for Birds 25 and 22W for each condition according to the following equation: $[(\text{pellets earned} - \text{min})/(\text{max} - \text{min})] \times 100$. A value of zero indicates that the bird earned the minimum number of pellets in that condition; a value of 100 indicates that the bird earned the maximum number of pellets in that condition. The results for each bird for each condition are presented in Figure 3; the mean of all replications of a given condition is shown. The horizontal lines indicate the percentage of pellets that would be earned if the pigeon were to respond randomly. Against all three of the TFT conditions (TFT, TFT + 1, TFT × 2), the pigeons earned little more than the minimum number of pellets; Bird 25 received an average of only 5.7%, and Bird 22W earned only 10% of the difference, even less than the 62.5% they would have earned had they responded randomly. This contrasts with the 78% to 99% earned when playing against the TFT-dom, chicken, and RND strategies, where earned pellets approached the maximum.

EXPERIMENT 3

A final experiment extended the generality and robustness of the previous results. Over the 10 TFT conditions studied in Experiment 1, there was only one case of cooperation; over the 14 TFT conditions in Experiment 2, there was no case of cooperation. Despite these facts, the small number of subjects studied might be cause for wariness. Consequently, 7 additional birds were studied against two different computer strategies in the final experiment. Both TFT and TFT-dom were examined to see whether defection would be obtained reliably under the former and cooperation would be obtained under the latter.

METHOD

Subjects and Apparatus

Seven male White Carneau pigeons, with previous experience pecking keys for mixed

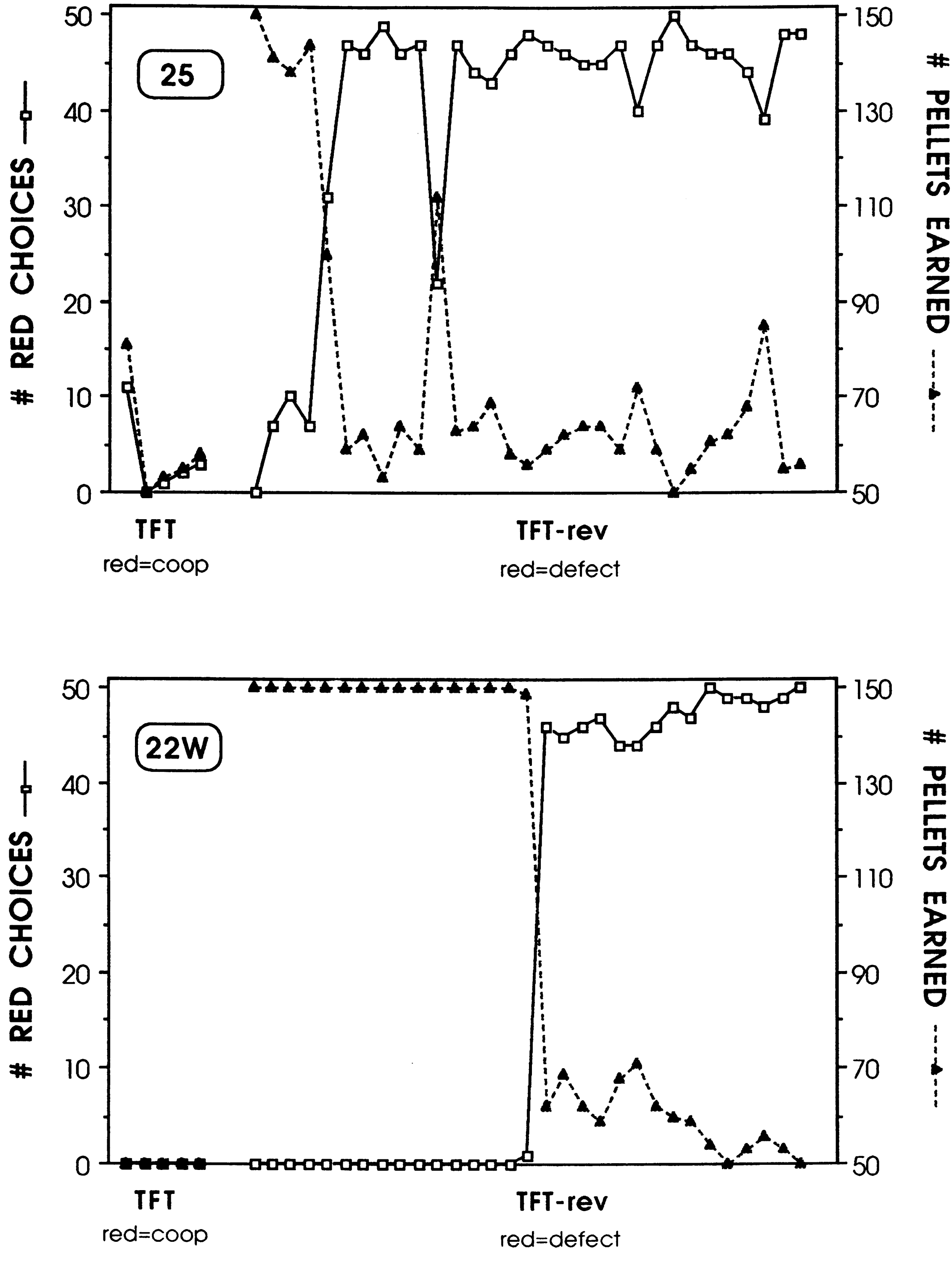


Fig. 2. Number of red choices and pellets earned for the final five sessions of TFT and the 31 sessions of TFT-rev (Conditions 10 and 11, respectively, of Experiment 1) for Birds 25 and 22W. During TFT, red meant cooperate; during TFT-rev, red meant defect.

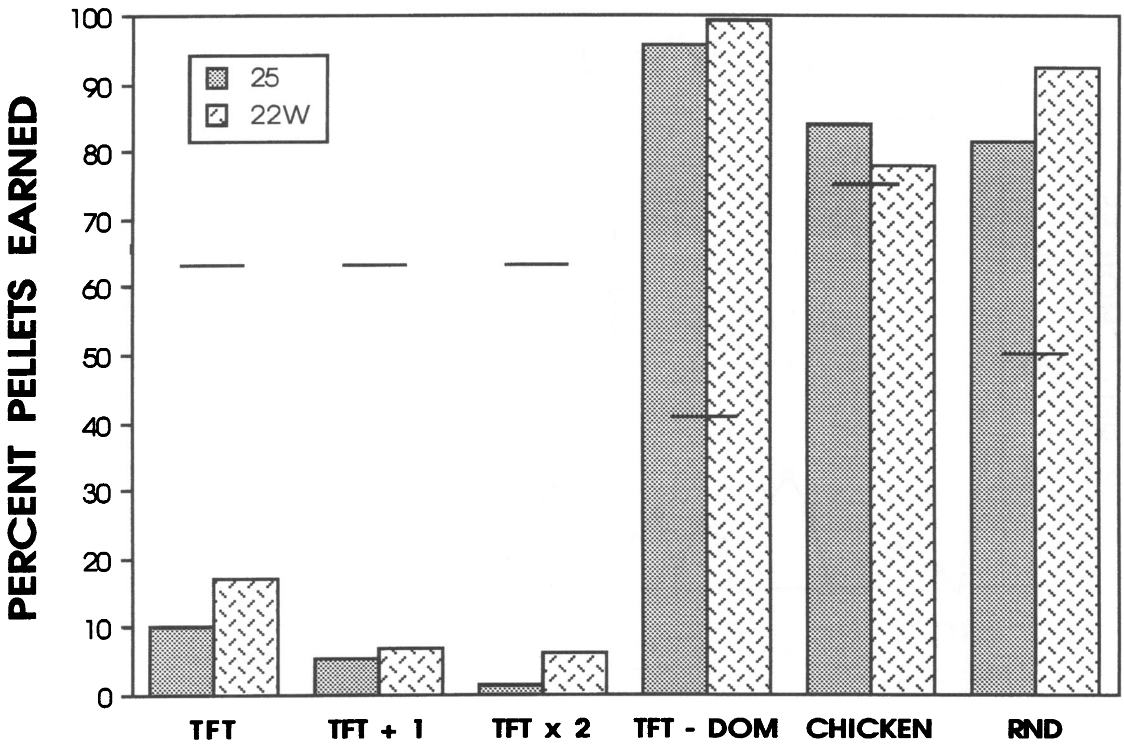


Fig. 3. Mean percentage of the difference between the minimum and the maximum number of food pellets possible that was earned by Birds 25 and 22W in each condition.

grains, were maintained as in the prior experiments. The apparatus was the same as that used previously.

Procedure

The pigeons were first trained to peck the red and green keylights for food pellets by an autoshaping procedure and were then placed on the first of three experimental conditions. All details of the procedure were the same as those in Experiments 1 and 2. The red keylight and blue cuelight represented cooperation by the pigeon and computer, respectively, and the green keylight and yellow cuelight represented defection by the pigeon and the computer, respectively. Each bird was studied against both TFT and TFT-dom, with the payoffs as noted in Table 1.

For 3 of the birds (15, 16, and 17) the first experimental condition was TFT. After 15 sessions, the condition was changed to TFT-dom for 15 sessions, followed by a return to TFT for 45 sessions. The other 4 birds (6, 7, 9, 10) were studied in the reverse order: TFT-dom followed by TFT and then a return to TFT-dom. TFT-dom was studied for 15 sessions (30 for Bird 9 on its replication) and TFT was studied for 30 or 60 sessions.

Results and Discussion

The median number of cooperations (red key choices) from the last 5 days is shown in Figure 4 for each bird for all conditions in the order in which they were studied. It is clear that defection (green key choices) was the consistent steady-state response when playing against the TFT conditions, whereas cooperation occurred against TFT-dom. The difference in the median number of cooperations during the birds' first exposure to TFT and during their first exposure to TFT-dom was statistically significant ($t = 19.34$, $p < .001$), as was the difference in number of pellets earned ($t = 13.31$, $p < .001$).

The mean number of cooperations averaged across all 7 birds under all 10 TFT conditions (based on the medians of the last five sessions of each condition) was 4.5 (out of a possible 50) whereas for the 11 TFT-dom conditions, the mean number of cooperations

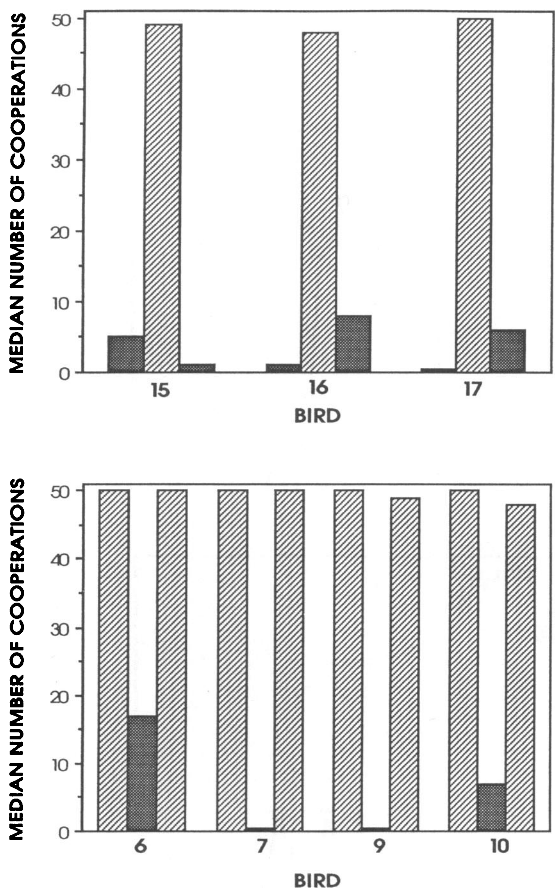


Fig. 4. Median number of cooperations from the final five sessions under TFT and TFT-dom, in the order in which these conditions were studied for each bird in Experiment 3. Dark bars represent results from TFT conditions; lighter bars represent results from TFT-dom conditions.

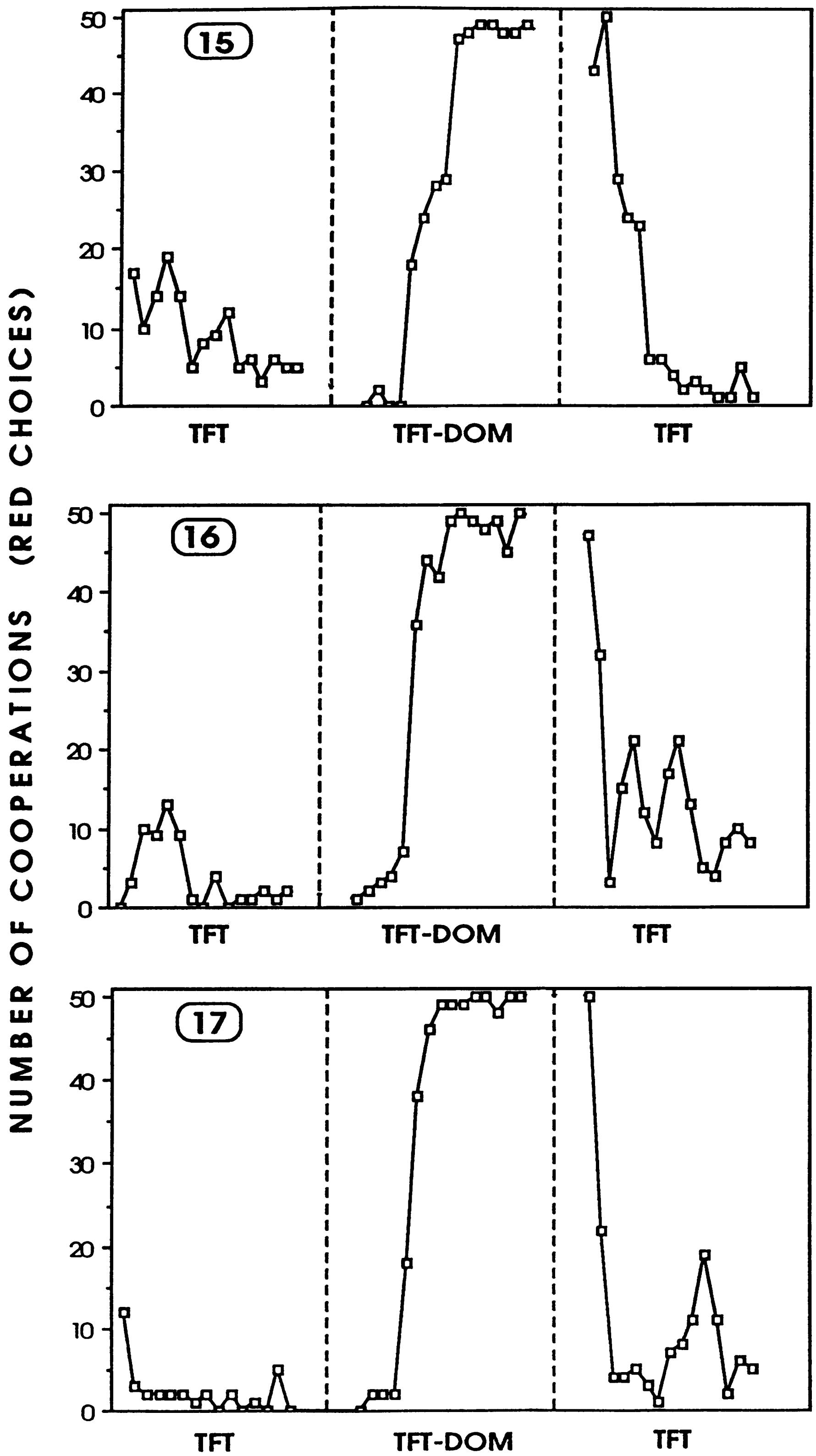
was 49.5. The corresponding mean number of pellets earned under TFT was 62.9 (the minimum possible being 50 and the maximum being 150); under TFT-dom, the mean number of pellets earned was 148.2 (with the minimum being 25 and the maximum being 150). The percentage of the difference between the minimum and the maximum number of pellets possible that was actually earned averaged 12.9% and 98.5% against TFT and TFT-dom, respectively.

Session-by-session results are presented for Birds 15, 16, and 17 in Figure 5. Plotted are the number of cooperations (choices of the red key) for each of the 15 days of the TFT and the TFT-dom conditions. For the replications of the TFT condition, the medians of 3-day blocks are plotted because this replication was studied for 45 days. Notice that all 3 birds came to defect (choose green) during the initial TFT condition and earned only 64, 56, and 50 pellets, respectively, on the 15th session. When the computer strategy was changed to TFT-dom, each bird switched to pecking the red (cooperation) key, almost tripling the number of pellets earned. This change was apparent within 5 or 6 days. Interestingly, when the computer strategy was then returned to TFT, the birds initially continued pecking red and thus continued to earn close to the maximum number of pellets. Indeed, for the first 2 days, Birds 15, 16, and 17 cooperated an average of 45, 49, and 50 trials and earned 145, 149, and 149 pellets. However, with further exposure to the TFT condition, each bird shifted its choice to the green (defect) key, and the number of pellets earned dropped from near maximum to close to the minimum (53, 67, and 68 pellets on the final session of TFT). This pattern of behavior replicated that of the previous experiments.

These results thus support further the conclusions from Experiments 1 and 2: When playing against TFT, the birds defected. However, the birds were sensitive to the pay-off structure and altered their responses accordingly: When TFT-dom was instituted, the birds altered their choice to one of cooperating. Once again, when cooperation earned the greatest immediate as well as overall total reward, as was true under TFT-dom, the birds cooperated. However, when cooperation earned the greatest overall reward but defection earned the greatest immediate reward, as was the case under TFT, the birds defected.

GENERAL DISCUSSION

In general, the pigeons' choice responding on 2 × 2 games was controlled by the reinforcer outcomes of the immediately occurring trial, with the pigeons choosing the outcome with the immediately higher pay-off. The birds maximized overall reward only when such behavior was consistent with maximizing reward on the immediately occurring trial. When maximizing overall reward was inconsistent with maximizing reward on the immediate trial, the birds responded to the im-


NUMBER OF COOPERATIONS (RED CHOICES)
15
16
17
0
10
20
30
40
50
TFT
TFT-DOM
TFT

mediate reward situation. Such behavior considerably reduced total pellets earned under certain conditions. The birds cooperated only when such behavior achieved the highest immediate reward (e.g., against TFT-dom).

Although it would have been in the pigeons' long-term best interest always to cooperate when playing the prisoner's dilemma game against an opponent using a tit-for-tat strategy, they failed to do so. Instead, by almost always defecting, the birds earned only about one third of the number of food pellets that they could have earned if they had cooperated. To expect any player always to cooperate when playing a prisoner's dilemma against an opponent using a tit-for-tat strategy, it must be assumed that the player integrates over several trials of the iterated prisoner's dilemma game. The pigeons, perhaps because of their steep temporal discount functions (Green, Fisher, Perlow, & Sherman, 1981; Kagel & Green, 1987), did not do this. Thus, they made the dominant response, to defect, on each trial.

The effects of temporal discounting on rates of cooperation in the iterated prisoner's dilemma game have been discussed by Axelrod (1984). Consider a player who would earn a pay-off of three on each of some indefinite number of trials by always cooperating against an opponent playing tit for tat. On Trial 1, when the player must decide which choice to make, the utility from cooperating is 3. The utility from cooperating on Trial 2, however, is only a fraction of what it is on Trial 1, and the utility from cooperating on Trial 3 is a fraction of what it is on Trial 2, and so on. The weight (or importance) of a later move relative to an earlier move is called w, a discount parameter, which "represents the degree to which the payoff of each move is discounted relative to the previous move" (Axelrod, 1984, p. 13). The larger the value of w, the more the pay-offs from later trials are taken into account in the calculation of total pay-offs.

The total utility from always cooperating for the player in question would be $3 + 3w + 3w^2 + 3w^3 \ldots$. The utility from always defecting when the opponent is playing TFT, assuming $T = 5$ and $P = 1$, would be $5 + w + w^2 + w^3 \ldots$. Axelrod (1984) notes that the sum of this infinite series for $0 < w < 1$ would be $3/(1 - w)$ and $5 + [w/(1 - w)]$ for the two cases, respectively. For example, if w were large, say .9, then the utility from always cooperating would be 30 [i.e., $3/(1 - .9)$], whereas that from defecting would be 14 ($5 + [.9/(1 - .9)]$). On the other hand, if w were small, say .1, then the utility from cooperating would be 3.3, whereas that from defecting would be 5.1. In general, then, always cooperating maximizes a player's utility only if w is sufficiently large; otherwise defection is the optimal choice. The steep discounting by pigeons of future outcomes (Green et al., 1981; Kagel & Green, 1987) suggests that in the present study w was too small to promote cooperation in the pigeons playing against tit for tat.

In a related vein, the pigeons' behavior can be seen as exhibiting a lack of self-control. Self-control may be defined as the choosing of a larger but delayed reward over a smaller but more immediate reward. Lack of self-control, impulsiveness, is choosing the smaller but more immediate reward over the larger but delayed reward (Green, 1982; Logue, 1988). More generally, self-control can be conceptualized as responding to the overall molar reward contingencies, whereas impulsiveness is responding to the more immediate molecular reward contingencies. When faced with a choice between a small reward at time t and a larger reward at time $t + x$, pigeons choose the small reward when the value of t is small. Their preference reverses to that of the larger reward as the time t is increased (Green et al., 1981). In fact, pigeons will come to make a commitment response when the value of t is large in order to insure the receipt of the larger reward. Such a commitment response restricts future choice so that as time to the rewards approaches, the pigeon cannot act on its reversal in preference and select the smaller but more immediate

←

Fig. 5. Number of cooperations (red-key choices) for each session of TFT, TFT-dom, and (in blocks of 3-day medians) the replication of TFT for Birds 15, 16, and 17.

reward but now must receive the larger reward (Rachlin & Green, 1972).

The pigeons in the present study faced similar choice options against TFT. Following a defection, the bird could cooperate and receive a larger reward (three pellets), but only following a trial in which it must first receive the zero-pellet outcome. On the other hand, it could continue to defect and thus receive a small reward (one pellet) on that trial. Thus, the birds generally chose the smaller but more immediate (one-pellet) reward to that of the larger but delayed (three-pellet) reward. Similarly, following a cooperation, the bird could earn five pellets on that next trial followed by one pellet on the ensuing trials were it to defect repeatedly; on the other hand, it could earn three pellets on all of the ensuing trials were it to continue to cooperate. The birds generally opted for the immediate benefit of five pellets (which would then likely lead to future one-pellet rewards) rather than to respond to the molar reward contingencies and cooperate, which would lead to future three-pellet rewards. In the prisoner's dilemma game as structured here, the bird could not commit itself to the long-term, higher valued outcome; it consequently defected and thus earned an overall lower valued but more immediate outcome.

The present results may have implications for the study of cooperation and competition in both nonhuman animals and in humans. Any implications and generalizations, however, must be tempered in light of the fact that as an initial foray into the field, the present work examined behavior under a limited set of matrix values and other parameters. Even though the pay-offs were varied over a reasonably meaningful range, still greater variations in relative and absolute size of the payoffs are in order. It also is to be noted that only positive consequences were used. The effects of a mixed-outcome matrix remain to bc evaluated. With such caveats in mind, however, we suggest that the ability to learn to respond optimally in a generalized prisoner's dilemma game may be rare among certain animals because of the problems of temporal discounting and self-control. This is not to say, of course, that animals would not come to cooperate or cannot learn a tit-for-tat strategy or that it has not evolved under certain well-defined domains. Clearly, the ability to learn from experience is also a product of evolution (Skinner, 1966; Staddon, 1988), but unless learning mechanisms have evolved to be quite sensitive to future outcomes, the ontogenetic development of cooperation under a prisoner's dilemma situation would remain a rarity in nonhuman animals that discount delayed events so extensively. Indeed, discounting future rewards may well be an adaptive response to uncertainty in an animal's natural environment (Kagel, Green, & Caraco, 1986).

How, then, do our results relate to the issue (raised in the Introduction) of how cooperative behavior could have evolved among populations of organisms that were initially noncooperative? It has been proposed that if members of a population were to appear who, for whatever reason, play a tit-for-tat strategy, then cooperation would evolve in the population. If individuals were to play TFT, they would earn greater pay-offs when playing against each other than would those who don't play TFT when they play against each other. In addition, when those playing TFT play against those who don't, the TFT players are no worse off than the others. Consequently, TFT offers a net advantage, and once it enters the population, no individual with a different pattern of behavior could do as well.

The present work may serve as a benchmark for assessing this line of theorizing. The computer was programmed to play TFT and in effect served as a probe for examining the behavior that emerged in the pigeons. Granting the general assumptions of theories that attempt to account for the evolution of cooperation (e.g., that interactions between individuals in the population have a sufficiently large probability of continuing; Axelrod & Hamilton, 1981) and mindful of the caveats noted earlier, the present results suggest that cooperation does not evolve under circumstances such as those in the present study. Of course, cooperative behavior might evolve when rate of temporal discounting is considerably lower, commitment responses are permitted, or delays between choice and outcome are sufficiently large. These possibilities remain to be tested.

Humans are clearly more sensitive to future outcomes than are pigeons, showing shallower discount functions (Green, Fry, &

Myerson, 1994; Rachlin, Raineri, & Cross, 1991). Outside of the laboratory, however, intertrial times in a prisoner's dilemma-like situation may be years, decades, or generations, rather than minutes. When such long time spans are involved, temporal discounting may be responsible in part for phenomena such as continued defection by both players in an arms race (see Snyder & Diesing, 1977) and the general refusal among industrialized nations to conserve natural resources. The need to find ways of bringing long-term consequences to bear on our immediate behavior must be met if we are to live cooperatively with each other and with our environment.

REFERENCES

Axelrod, R. (1980a). Effective choice in the prisoner's dilemma. *Journal of Conflict Resolution, 24,* 3–25.

Axelrod, R. (1980b). More effective choice in the prisoner's dilemma. *Journal of Conflict Resolution, 24,* 379–403.

Axelrod, R. (1984). *The evolution of cooperation.* New York: Basic Books.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211,* 1390–1396.

Baefsky, P. M., & Berger, S. E. (1974). Self-sacrifice, cooperation and aggression in women of varying sex-role orientations. *Personality and Social Psychology Bulletin, 1,* 296–298.

Baum, A., & Gatchel, R. J. (1981). Cognitive determinants of reaction to uncontrollable events: Development of reactance and learned helplessness. *Journal of Personality and Social Psychology, 40,* 1078–1089.

Boyd, R., & Lorberbaum, J. P. (1987). No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature, 327,* 58–59.

Dugatkin, L. A. (1988). Do guppies play tit for tat during predator inspection visits? *Behavioral Ecology and Sociobiology, 23,* 395–399.

Flood, M., Lendenmann, K., & Rapoport, A. (1983). 2 × 2 games played by rats: Different delays of reinforcement as payoffs. *Behavioral Science, 28,* 65–78.

Gardner, R. M., Corbin, T. L., Beltramo, J. S., & Nickell, G. S. (1984). The prisoner's dilemma game and cooperation in the rat. *Psychological Reports, 55,* 687–696.

Green, L. (1982). Self-control behavior in animals. In V. L. Smith (Ed.), *Research in experimental economics* (Vol. 2, pp. 129–150). Greenwich, CT: JAI Press.

Green, L., Fisher, E. B., Jr., Perlow, S., & Sherman, L. (1981). Preference reversal and self-control: Choice as a function of reward amount and delay. *Behaviour Analysis Letters, 1,* 43–51.

Green, L., Fry, A. F., & Myerson, J. (1994). Discounting of delayed rewards: A life-span comparison. *Psychological Science, 5,* 33–36.

Insko, C. A., Schopler, J., Drigotas, S. M., Graetz, K. A., Kennedy, J., Cox, C., & Bornstein, G. (1993). The role of communication in interindividual-intergroup discontinuity. *Journal of Conflict Resolution, 37,* 108–138.

Kagel, J. H., & Green, L. (1987). Intertemporal choice behavior: Evaluation of economic and psychological models. In L. Green & J. H. Kagel (Eds.), *Advances in behavioral economics* (Vol. 1, pp. 166–184). Norwood, NJ: Ablex.

Kagel, J. H., Green, L., & Caraco, T. (1986). When foragers discount the future: Constraint or adaptation? *Animal Behaviour, 34,* 271–283.

Lazarus, J., & Metcalfe, N. B. (1990). Tit-for-tat cooperation in sticklebacks: A critique of Milinski. *Animal Behaviour, 39,* 987–988.

Logue, A. W. (1988). Research on self-control: An integrating framework. *Behavioral and Brain Sciences, 11,* 665–679.

Lombardo, M. P. (1985). Mutual restraint in tree swallows: A test of the tit for tat model of reciprocity. *Science, 227,* 1363–1365.

Masters, W. M., & Waite, T. A. (1990). Tit-for-tat during predator inspections, or shoaling? *Animal Behaviour, 39,* 603–604.

Maynard Smith, J. (1982). *Evolution and the theory of games.* Cambridge, England: Cambridge University Press.

Milinski, M. (1987). Tit for tat in sticklebacks and the evolution of cooperation. *Nature, 325,* 433–435.

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature, 364,* 56–58.

Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior, 17,* 15–22.

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior, 55,* 233–244.

Rapoport, A., Guyer, M. J., & Gordon, D. G. (1976). *The 2 × 2 game.* Ann Arbor: The University of Michigan Press.

Selten, R., & Hammerstein, P. (1984). Gaps in Harley's argument on evolutionary stable learning rules and in the logic of "tit for tat." *Behavioral and Brain Sciences, 7,* 115–116.

Skinner, B. F. (1966). The phylogeny and ontogeny of behavior. *Science, 153,* 1205–1213.

Snyder, G. H., & Diesing, P. (1977). *Conflict among nations: Bargaining and decision making in international crises.* Princeton, NJ: Princeton University Press.

Staddon, J. E. R. (1988). Learning as inference. In R. C. Bolles & M. D. Beecher (Eds.), *Evolution and learning* (pp. 59–77). Hillsdale, NJ: Erlbaum.

Received January 26, 1993
Final acceptance March 1, 1995