RESPONSE TO PARTIAL TIT-FOR-TAT STRATEGIES IN THE PRISONER'S DILEMMA GAME

Jonathan Pool
Department of Political Science, DO-30
University of Washington
Seattle, Washington 98195, USA

and

Bernard Grofman
School of Social Science
University of California, Irvine
Irvine, California 92717, USA

ABSTRACT

Response to Partial Tit-for-Tat Strategies in the Prisoner's Dilemma Game

An experiment was conducted to test whether players of iterated Prisoner's Dilemma games would respond to small variations in the probability with which their opponent reciprocated their latest move, as an earlier analysis showed they would if they behaved optimizingly. Forty-one West German adolescent subjects played 3 games of 100 rounds each against a computer programmed with 2 different probabilistic tit-for-tat strategies, conducive to 2 different rational counter-strategies. Subjects received either no information or a hint about the possible nature of the computer's strategy.

Comparisons of the rate of cooperation among subjects facing the 2 different strategies showed that both within- and between-subject behavior varied as predicted on the basis of rationality assumptions. The variation in cooperation was observed at both levels of information about the computer's strategy. High information, however, had the unexpected effect of increasing the tendency to cooperate, even when cooperation was counter-optimal.

Recent theoretical work has proven that small variations in the strategy of one player in a Prisoner's Dilemma game can elicit large (in fact, total) fluctuations in the behavior of the other player, if the latter is playing rationally. This study is a first attempt to discover whether such variations will produce a rational response in real players.

A symmetric Prisoner's Dilemma game, the most studied variety of 2-person game, is defined by a payoff matrix

		Column 1	Column 2
Row	1	a,a	b,c
	2	c,b	d,d

where each pair of entries indicates the payoffs to players Row and Column, respectively, if they make the combination of moves indicated in the margins of the matrix, and where the values of the entries are constrained by two inequalities:

(1) $c > a > d > b$

(2) $2a > b + c > 2d$.

Moves 1 and 2 are conventionally called "cooperation" and "defection", respectively.

Despite some analysts' ingenious efforts to minimize this conclusion, defection is always the optimizing move in one-shot Prisoner's Dilemma (PD) games without coordination between the players (identifying reference 1). When PD games are iterated, however, it is possible to find a strategy that will induce one's opponent, if self-interested and rational, to cooperate on every move but the last (identifying reference 2; identifying reference 3). Specifically, there are partial tit-for-tat strategies that will have this effect. These are strategies according to which each move has a certain probability, p, of being identical to the other player's previous move.

For payoff matrices satisfying the common condition, $a + d = b + c$, we showed that an optimizing player who sees that (s)he is playing against a partial tit-for-tat strategy and that the opponent's strategy is not subject to influence will find it rational either to cooperate on every move or to defect on every move. To determine whether constant cooperation or constant defection is optimal, one need merely compare the p of one's opponent with p_0, where

$$p_0 = \frac{c - b}{a - b + c - d} = \frac{c - b}{2(a - b)}.$$

If the opponent is copying one's last move more often than p_0, i.e. if $p > p_0$, constant cooperation is optimal; if $p < p_0$, constant defection is optimal; and if $p = p_0$, all strategies have the same effect. Furthermore, the greater is p the more is won (or the less lost) on the average by a strategy of pure cooperation.

In spite of the attention given to the conditions of cooperation and defection in PD games, the experimental literature reveals hardly anything about whether tit-for-tat (TFT) strategies differing in completeness (i.e., p) elicit different frequencies of cooperation. This is evident in Oskamp's (1971) comprehensive review, and we are not aware of any studies of this question since its publication. The few studies that have manipulated p have made its values grossly different (e.g., 1 and .5), and cooperation has been greater with the larger p, as expected. But even minor differences in p, as long as they straddle p_0 for the payoff matrix being used, will alter the preferability of the two available moves and hence should affect the frequency of cooperation. If p changes from below to above p_0 or vice versa, cooperation should rise or fall, respectively. One can also expect that such changes will be magnified by information that helps players discover that their opponents are using a TFT strategy.

METHOD

Subjects

Subjects were 41 persons, aged 15 to 21, who lived in the area of Mannheim and Heidelberg, Federal Republic of Germany. They had previously taken part in a social-psychological experiment and indicated a desire to be subjects again. (Some of those previously classified as highly introverted or extraverted had been withdrawn from the pool for other purposes.) Most of the subjects were secondary school students. A ratio of 25 males to 16 females was attained by accepting all female volunteers but not all male ones. Subjects underwent the experiment in groups of 6 to 13 over a period of 5 days.

Apparatus

The experiment took place in the computer-assisted instruction laboratory of a large training institution. Each subject was assigned to an IBM 2740 typewriter terminal inside a 3-walled carrel, connected to an IBM 370/155 computer, which executed the randomizations, instruction giving, training, move recording, opponent simulation, result tabulation, feedback delivery, rule enforcement, and data analysis to be described below, via programs written in APL-PLUS.[1]

Procedure

Before the experiment, subjects were instructed in the use of the terminal. They were told that they were about to play 3 games against the computer, each lasting 100 rounds. On each round each player (i.e. the subject and the computer) would move "1" or "2" without knowing which move the other was making on that round. The payoff matrix was printed out and torn off by subjects for reference during the game. Translated into English, it appeared as follows:

IF YOU MOVE....	1	1	2	2	
AND IT MOVES...	1	2	1	2	
THEN YOU WIN...	5	-3	8	0	PFENNIGS

(Thus a = 5, b = -3, c = 8, d = 0.) After receiving these instructions, subjects had to complete satisfactorily a test of their understanding of how to make moves on the terminal and what the payoffs associated with combinations of moves were. Then the experiment began.

The computer began each round by displaying the round number. Subjects had as long as they wanted to make each move. After each move by a subject, the subject was informed of how the computer had moved and how much the subject had won or lost on that round. After rounds 5, 10, 20, 40, 60, 80, and 100, the subject was also given summary feedback on the cumulative results of the game to that point, showing in parallel the two players' moves since the beginning of the game, as well as the subject's total winnings and the winnings expressed as pfennigs per round. At the end of each game, the subjects' winnings were paid in cash by the experimenter.

Manipulations

Subjects were assigned randomly to one of 12 experimental conditions, consisting of all combinations of 4 strategy sequences and 3 information sequences. In any game, the computer followed either a $\frac{5}{8}$ or a $\frac{7}{8}$ TFT strategy. Approximately $\frac{1}{4}$ of the subjects faced each of the following sequences of computer strategies:

		Game		
		1st	2nd	3rd
Sequence	A	$\frac{5}{8}$	$\frac{7}{8}$	$\frac{5}{8}$
	B	$\frac{5}{8}$	$\frac{7}{8}$	$\frac{7}{8}$
	C	$\frac{7}{8}$	$\frac{5}{8}$	$\frac{5}{8}$
	D	$\frac{7}{8}$	$\frac{5}{8}$	$\frac{7}{8}$

The information sequences were determined by the amount of information the subject had about the computer's strategy in each game. There were 2 levels of information. At the low level, only the information described under "Procedure" above was given. At the high level, this information was supplemented with the following (in German):

IMPORTANT NEWS! A SPY IN THE COMPUTER ROOM REPORTS THAT THE COMPUTER FOLLOWS A PARTICULAR STRATEGY DURING THE GAME AND STAYS WITH IT UNTIL THE END OF THE GAME.

CAN YOU DISCOVER THIS STRATEGY AND THUS INCREASE YOUR WINNINGS FURTHER?

MORE IMPORTANT ADVICE!! IT IS POSSIBLE THAT THE COMPUTER'S STRATEGY DEPENDS ON YOUR PREVIOUS MOVES. IN THIS CASE, EACH OF YOUR MOVES COULD HAVE AN EFFECT ON LATER MOVES OF THE COMPUTER.

The three information sequences were:

		Game		
		1st	2nd	3rd
	X	low	low	low
Sequence	Y	low	low	high
	Z	high	high	high

The carrels and a prohibition on talking kept the low-information subjects from obtaining any of this advice from the high-information subjects.

RESULTS

Given the payoff matrix shown above,

$$p_0 = \frac{c - b}{2(a - b)} = \frac{8 + 3}{2(5 + 3)} = \frac{11}{16}.$$

Since $\frac{7}{8}$ is higher and $\frac{5}{8}$ is lower than p_0, a rational player would always cooperate in the $\frac{7}{8}$ and defect in the $\frac{5}{8}$ condition. Naturally, we do not expect to observe perfectly rational behavior, if only because of the inevitable error in inferring the computer's strategy from a finite sample of its moves. We do, however, expect to see more cooperation against a $\frac{7}{8}$ than against a $\frac{5}{8}$ TFT strategy.

If for exploratory purposes we collect the results from rounds 2 to 99 of all games in which the computer applied the same strategy, ignoring temporarily the strategy and information sequences, we find this expectation confirmed (see Table 1).[2] Less than a third of the moves against the $\frac{5}{8}$ strategy were cooperative, compared with almost half of the moves against the $\frac{7}{8}$ strategy. Of the 41 subjects, all but 6 cooperated on a larger proportion of these rounds in the $\frac{7}{8}$ than in the $\frac{5}{8}$ TFT games (sign test: $p < .00001$). The results in Table 1 further suggest a major asymmetry in subjects' tendencies toward cooperation and defection. That is to say, when defection is optimal ($\frac{5}{8}$ condition), actual behavior is strongly prone to defection. But when cooperation is optimal ($\frac{7}{8}$ condition), subjects' behavior is still prone to defection, though less so. This is true in spite of the fact that the excess of $\frac{7}{8}$ over p_0 is greater than the shortage of $\frac{5}{8}$ under p_0.[3] In short, a $\frac{7}{8}$ TFT rewards cooperation more than a $\frac{5}{8}$ TFT rewards defection, yet a $\frac{7}{8}$ TFT produces cooperation less often than a

$\frac{5}{8}$ TFT produces defection.

Having explored the major features of the results with collapsed conditions, we can now look more carefully at the impact of the different levels of p. During the first 2 games, with the information level held constant for each subject, half the subjects were shifted from a $p = \frac{5}{8}$ to a $p = \frac{7}{8}$ condition and the other half from a $\frac{7}{8}$ to a $\frac{5}{8}$ computer strategy (see "Manipulations" above). This provides an opportunity to test the hypothesis that subjects will increase their rate of cooperation when the opponent increases the p of its TFT strategy, and will decrease cooperation if the opponent's p goes down. Table 2 shows that this indeed happens. Cooperation among subjects facing a rising p increased overall from under a third to nearly a half. Sixteen of the 21 subjects in this condition raised their cooperation rates from game 1 to game 2 (sign test: $p < .02$). Subjects facing a falling p cooperated 42.4% of the time in the first game but only 25.8% of the time during game 2. Seventeen out of these 20 subjects lowered their cooperation rates (sign test: $p < .002$).

The level of information possessed by subjects was not expected to reverse the relationships hypothesized above. At either of the 2 information levels, subjects facing the $\frac{7}{8}$ TFT strategy were expected to cooperate more often than those facing the $\frac{5}{8}$ strategy. Likewise, within each information level, the change in cooperation from the first to the second game was expected to be in the same direction as the change in the computer's p. The effect that was hypothesized was that high information would generally increase the frequency of optimizing behavior, thus strengthening the already discussed relationships.

In fact, the effects of information were partly in the opposite direction from what was expected. One expectation that was confirmed was that the $\frac{7}{8}$ TFT strategy would elicit more cooperation than the $\frac{5}{8}$ strategy when information was controlled for. We can control for information and eliminate any sequence effects by confining our attention to game 1, and, as Table 3 shows, the rate of cooperation was higher against a $\frac{7}{8}$ TFT than against a $\frac{5}{8}$ TFT, within each information condition in that game. In a 2-way analysis of variance, this main effect was significant ($F_{(1,37)} = 4.10$, $p = .05$).

The expected residual interaction between information and strategy conditions was evident in game 1, but it was not very significant ($F_{(1,37)} = 3.07$, $p < .09$), and the level of information made hardly any difference in the mean cooperation rate against the $\frac{5}{8}$ TFT strategy.

For our last hypothesis to be convincingly supported, the changes from game 1 to game 2 would have had to be greater in the direction of optimality under high information. Cooperation would have had to increase more sharply as the computer shifted from $\frac{5}{8}$ to $\frac{7}{8}$, and decrease more sharply as it shifted from $\frac{7}{8}$ to $\frac{5}{8}$, for the high-information subjects than it did in the low-information condition. Yet this difference obtained only when increasing cooperation was optimal ($\frac{5}{8}$ to $\frac{7}{8}$ condition). When decreasing cooperation was optimal ($\frac{7}{8}$ to $\frac{5}{8}$ condition), high information led to less of a decrease than did low information. Thus the expected residual interaction effect between information and strategy conditions on the increase of cooperation between game 1 and game 2 did not materialize ($F_{(1,37)} = .14$, $p > .7$). Instead, high information had a positive effect on the rate of increased cooperation ($F_{(1,37)} = 2.98$, $p < .10$). It eliminated much of the leaning toward defection that we noted above, even when this leaning was beneficial. Where the most profitable course was to stop cooperating, ignorance turned out to be bliss.[4] Table 4, bringing the results of the 3 games together, shows that, in all the games where defection was the optimal response, low-information subjects moved optimally in $\frac{3}{4}$ of all rounds, while high-information subjects moved optimally less than $\frac{2}{3}$ of the time.

CONCLUSION

The principal finding of this study is that differences and changes in the strategy of one player in a Prisoner's Dilemma game--though they may seem small--bring about differences and changes in the behavior of the other player, not only in rational theory but also in fact. Cooperation is more common when it is the optimizing strategy. When the optimizing strategy changes, strategic behavior changes in the same direction. This is true regardless of whether the ideally optimizing behavior changes from cooperation to defection or vice versa, and regardless of whether the subject has low or high information about the kind of strategy the opponent is using. This means that, as a first approximation, rationality assumptions applied to PD games turn out to be predictive of actual behavior.

What made it possible to use a calculus of rationality in the first place was the fact that in this experiment subjects played against a programmed opponent. As we expect to confirm in studies currently underway, behavior in such games should be predictive of behavior in ordinary PD situations (with 2 human subjects) as well, as long as they include those devices commonly used in the real world to "program" oneself, i.e. convince the other player that one is irrevocably committed to one's own strategy. Our findings indicate that players who manage to commit themselves credibly to a partial TFT strategy ($p > p_0$) will eventually reap the benefit of eliciting more cooperation in their opponents than they practice themselves.[5] How players discover and implement such a strategy remains to be explored.

TABLE 1

Cooperation Rates against Different Computer Strategies

Round(s)	Computer TFT Strategy		
	5/8	7/8	Both
1	.468 (62)	.475 (61)	.472 (123)
2-99	.303 (6076)	.485 (5978)	.393 (12054)
100	.081 (62)	.197 (61)	.138 (123)
All	.302 (6200)	.482 (6100)	.392 (12300)

Proportion is based on total number of moves in parentheses.

TABLE 2

Cooperation Rates against Changing Computer Strategies

Computer's Change of Strategy from 1st to 2nd Game	Game			N
	1st	2nd	Both	
5/8 to 7/8	.315 (2058)	.476 (2058)	.395 (4116)	21
7/8 to 5/8	.422 (1960)	.258 (1960)	.340 (3920)	20
Both	.367 (4018)	.369 (4018)	.368 (8036)	41

Proportion is based on total number of moves in parentheses.

Moves 2-99 of each game are included.

TABLE 3

Cooperation Rates against Changing Computer Strategies with Different Information Levels

Information about Computer's Strategy	Computer's Change of Strategy from 1st to 2nd game	Game			N
		1st	2nd	Both	
Low	5/8 to 7/8	.322 (1274)	.453 (1274)	.387 (2548)	13
	7/8 to 5/8	.365 (1372)	.165 (1372)	.265 (2744)	14
	Both	.344 (2646)	.303 (2646)	.324 (5292)	27
High	5/8 to 7/8	.304 (784)	.513 (784)	.408 (1568)	8
	7/8 to 5/8	.554 (588)	.474 (588)	.514 (1176)	6
	Both	.411 (1372)	.496 (1372)	.454 (2744)	14

Proportion is based on total number of moves in parentheses. Moves 2-99 of each game are included.

TABLE 4

Cooperation Rates against Different Computer Strategies with Different Information Levels

Information about Computer's Strategy	Computer TFT Strategy		
	5/8	7/8	Both
Low	.252 (3332)	.428 (3234)	.339 (6566)
High	.365 (2744)	.553 (2744)	.459 (5488)
Both	.303 (6076)	.485 (5978)	.393 (12054)

Proportion is based on total number of moves in parentheses. Moves 2-99 of all games are included.

FOOTNOTES

2. Since moves on round 1 were made before any evidence about the computer's strategy became available to the subject, there is no reason to expect any difference between the rate of cooperation against the two strategies on move 1. Since last moves, as stated earlier, are not governed by the considerations of an iterated PD game but rather can be considered as one-shot games, where defection is always optimal, there is no reason to expect the cooperation rate on last moves to be influenced by the computer's strategy; cooperation should, however, be lower on last moves than otherwise. Hence the subsequent analysis is based on rounds 2-99. Table 1 indicates that essentially the same proportion of subjects cooperated on round 1 regardless of strategy, as expected, and that 100th-round cooperation was very infrequent, as expected. Last-move behavior does appear, however, to have been influenced by the computer's strategy just as if the game had been going to continue. This phenomenon deserves further investigation.

3. The difference between $\frac{7}{8}$ and p_o is greater in three ways: absolutely ($\frac{3}{16}$ vs. $\frac{1}{16}$), proportionally (60% of the distance to 1 vs. 33% of the distance to .5), and in terms of the differences in payoffs between optimal and counter-optimal play (3 pf. per round vs. 1 pf. per round).

4. The possibility that in the $\frac{5}{8}$ TFT condition this information may have misled some subjects into more cooperation while helping other subjects understand the benefits of less

cooperation is suggested by the nearly equal cooperation rates in game 1. If this happened, we would expect a greater variance of cooperation levels in the high-information than in the low-information condition against the $\frac{5}{8}$ strategy. And indeed the variances in total number of cooperative moves are 282.5 and 134.8, respectively.

5. Bixenstine, Chambers, & Wilson (1964) point out that this takes place whenever the opponent's cooperation rate exceeds .5.

REFERENCES

Bixenstine, V.E., Chambers, N., & Wilson, K.V. Effect of asymmetry in payoff on behavior in a two-person non-zero-sum game. *Journal of Conflict Resolution*, 1964, *8*, 151-159.

Oskamp, S. Effects of programmed strategies on cooperation in the Prisoner's Dilemma and other mixed-motive games. *Journal of Conflict Resolution*, 1971, *15*, 225-259.

IDENTIFYING FOOTNOTE

1. Support for the programming, program revision, and conduct of this experiment was provided by Stiftung Rehabilitation (Heidelberg), the Deutscher Akademischer Austauschdienst, the Council for European Studies, the Institut für Sozialwissenschaften der Universität Mannheim, the State University of New York, and Sonderforschungsbereich 24 der Universität Mannheim. Assistance of various kinds was rendered by Rudolf Wildenmann, Betty Weneser, Walter Wehrli, Wilfried K. Schreiber, Margit Oswald, Martin Irle, Jeannette Huber, Dietrich Franz, Wolfgang E. Fendt, Günter Dhom, Walter Augsburger, and others.

IDENTIFYING REFERENCES

(1) Grofman, B. The Prisoner's Dilemma game and the problem of rational choice: paradox reconsidered. *Frontiers of Economics*, 1975, *1*, 101-119.

(2) Grofman, B., & Pool, J. Bayesian models for iterated Prisoner's Dilemma games. *General Systems*, 1975, *20*, 185-194.

(3) Grofman, B., & Pool, J. How to make cooperation the optimizing strategy in a two-person game. *Journal of Mathematical Sociology*, 1977, *5*, 173-186.

AUTHORS' NOTE

We are grateful for the detailed comments of Dale Hay on a previous draft.