Expecting Continued Play in Prisoner's Dilemma Games

A TEST OF SEVERAL MODELS

J. KEITH MURNIGHAN
Department of Business Administration
University of Illinois at Urbana—Champaign

ALVIN E. ROTH
Department of Economics
University of Pittsburgh

Several models of prisoner's dilemma interactions were tested in a series of twelve games whose termination point was determined probabilistically. A new model was introduced to discriminate among equilibrium and nonequilibrium situations on the basis of a player's expected benefits or losses for cooperating. The experiment included twelve payoff matrices, three probabilities for continuing, two opponent strategies, and the player's sex as independent variables. Results showed that both the game payoffs and the probability that the game would continue interacted to affect the rates of cooperation observed, and that the equilibrium model predicted this outcome most accurately. While the predictions of each of the models were supported, the equilibrium models appeared to be superior to the others. The discussion highlights the importance of considering the likelihood of a game terminating as a major determinant of the cooperation that can be expected in mixed—motive interactions.

This article presents an extension of a recent model of prisoner's dilemma games and reports an experiment investigating a new variable that may contribute to our understanding of the likelihood of individualistic behavior. The experiment examines the effects of the probability

AUTHORS' NOTE: The authors gratefully acknowledge the conscientious assistance of Kenneth Bettenhausen, who served as experimenter, and the constructive comments of S. S. Komorita. This research was partially supported by a grant from the National Science Foundation to the University of Illinois. Requests for reprints should be addressed to J. Keith Murnighan, Department of Business Administration, 104 Commerce West, 1206 South Sixth Street, Urbana, IL 61801.

JOURNAL OF CONFLICT RESOLUTION, Vol. 27 No. 2, June 1983 279-300

TABLE 1
A Representation of the Prisoner's Dilemma Game

		Player 2	
		Cooperate	*Defect*
Player 1	*Cooperate*	a, a	d, b
	Defect	b, d	c, c

NOTE: $b > a > c > d$

that an interaction will be repeated, along with two variables that have a long history in the research on prisoner's dilemma games: the payoffs obtainable by the players and the strategy of one's opponent. In addition, this study is the first test of the predictions of an equilibrium model of the prisoner's dilemma (Roth and Murnighan, 1978) with a large, diverse set of data.

The classical form of the prisoner's dilemma can be represented by the payoff matrices $A_1 = \begin{pmatrix} a & d \\ b & c \end{pmatrix}$ and $A_2 = \begin{pmatrix} a & d \\ b & c \end{pmatrix}$, where a, b, c, and d are numbers such that $b > a > c > d$.[1] Player 1 may choose a row of the matrices, while player 2 may choose a column. The payoff to player 1 is the element of A_1 thus specified, and the payoff to player 2 is the specified element of A_2 (see Table 1). When the payoffs conform to this ordering, the requirements for a prisoner's dilemma are fulfilled: Individualistic (defect) choices always yield higher payoffs than cooperative choices (game theoretically, individualistic choices *dominate* cooperative choices), but if both players make the individualistic choice, they each get a lower payoff than if both players had made the cooperative choice. For repeated play, an additional requirement that $b \; + < 2a$ is necessary. This makes it unprofitable for the players to establish a pattern of cooperative and individualistic choices.

SEVERAL INDICES OF COOPERATION

A variety of indices using the payoff values have been suggested as possible predictors of cooperation in prisoner's dilemma games. Rapo-

1. Much of the social psychological literature on prisoner's dilemma games has utilized the labels T (for Traitor), R (for Reward), P (for Punishment), and S (for Saint or Sucker) for the payoffs we have called b, a, c, and d. Because of additional notation used later in the article, the more neutral labels were used here.

TABLE 2
Summary of the Indices

Index		
r_1	=	a – c/b – d
r_2	=	a – d/b – d
r_3	=	c – d/b – d
r_4	=	b – a/b – d
e_1	=	b – a/a – c
e_2	=	b – a/a – d
g	=	$\frac{1}{1-p}[a - b + p(b - c)]$
k_1	=	a + d – b – c
k_2	=	a – d + b – c
k_3	=	a – d + b – c
k_4	=	a + d + b + c

port and Chammah (1965) proposed two indices for cooperation in symmetric games, r_1 and r_2 (see Table 2 for a summary of the indices and how they are calculated). While both emphasize that increases in a and decreases in b will contribute to greater benefits (and increased motivation) for mutual cooperation, r_2 places greater emphasis on the effects of the value of d. Indeed, by ignoring c, r_2 completely deemphasizes the mutual defect outcome as a motivator during play. Results reported by Steele and Tedeschi (1967) and Jones and associates (1968) both supported the predictions of r_1 and showed that different payoff matrices markedly affect the rates of cooperative choices.

Harris (1969) proposed two additional indices, r_3 and r_4. The r_3 index is a direct counterpart to Rapoport and Chammah's r_2 index: Where r_2 focused on the mutual cooperative payoff and ignored the mutual defect payoff, r_3 does the reverse. The r_4 index, on the other hand, focuses primarily on the maximum payoff that a player can possibly obtain. The "basis" or denominator for all four indices, the difference between b and d, also indicates a common agreement that the temptation offered by a very high payoff, relative to the lowest payoff in the matrix, underlies the likelihood of cooperative choice.

Research by Wyer (1969) was not strongly supportive of Harris's indices but, rather, supported three of the four indices (k_1 through k_4) suggested by Thibaut and Kelley's (1959) approach to social interaction (see Table 2). The first index represents the control one has over one's own outcomes. The second measures the "fate control" one player has over the other. The third measures the "behavior control" of one over the other. And the fourth index measures the overall level of outcomes available in the game. (All of these indices are written for symmetric games; they can be expanded easily for nonsymmetric games.)

Roth and Murnighan (1978) have proposed two additional indices, called equilibrium indices. They will be discussed below, along with a new related index presented here for the first time.

STRATEGIES

A second variable of concern in the literature on prisoner's dilemma games, and also a major independent variable in the present experiment, is the strategy used by one's opponent. Oskamp (1971) reviewed research concerning the effects of an opponent's choices on an individual's behavior in mixed-motive games. (Mixed-motive games are those in which elements of cooperation and competition can be present simultaneously.) He concluded that one of the most effective strategies leading to mutual cooperation was the *matching* strategy. A player using the matching strategy typically chooses cooperatively on the initial choice; subsequent choices duplicate one's opponent's previous choice. Thus, if one's opponent chooses CDCCDDC (C = cooperative choice, D = defect choice) over a seven-trial sequence, the matching strategy's response would be CCDCCDD. The matching strategy duplicates the opponent's choice with a one-trial lag.

More recently, Axelrod (1980a, 1980b) conducted two tournaments in which the matching strategy maximized a player's outcomes when paired, one at a time, with any member of a large set of opposing strategies. In their discussion of the matching strategy, Axelrod and Hamilton (1981) point out that the matching strategy possesses two particularly functional characteristics: It is contingent upon one's opponent's choices and it is forgiving. Thus, it does not allow one's opponent to take unfair advantage of altruistic or cooperative tendencies. At the same time, it holds a grudge for only the minimum amount of time. Indeed, the matching strategy seems to train an opponent into using the cooperative choice.

The current study investigated both the matching strategy and the *unforgiving* strategy. Like the matching strategy, the unforgiving strategy begins with a cooperative choice. It is also contingent on the opponent's responses because the unforgiving strategy continues to make the cooperative choice as long as the opponent does. However, once the opponent chooses the individualistic or defect choice, the unforgiving strategy chooses individualistically from that point on (with a one-trial lag). The CDCCDDC sequence facing an unforgiving strategist would result in a CCDDDDD response. Thus, where the matching strategy was contingent and forgiving, the unforgiving strategy is contingent but unforgiving.

THE PROBABILITY OF CONTINUING

Most research on the prisoner's dilemma game has involved repeated play, in one of two forms. The most prevalent has studied a long series (over 100 plays) of interactions between the same players, with a fixed endpoint to the game. The alternative, used in only two studies (Axelrod, 1980b; Roth and Murnighan, 1978), involved repeated play of the same game, with a fixed probability that the game would terminate. Roth and Murnighan (1978) focused on the values this probability takes as a determinant of the presence of cooperative equilibria within a game. An equilibrium results when both players are following strategies that give their opponent no incentive to adopt a different strategy. One equilibrium in the prisoner's dilemma game is for both players to defect at every trial.

If a game is to be played many times, consideration of the cooperative choice may increase, since players can attempt to influence their opponent's future behavior. However, the cooperative choice on the *last* trial has no possibility of influencing future play, and is thus always less profitable than the individualistic choice. There is no chance for retribution, and future gains are threatened. Thus, rational, profit-maximizing players will defect on the last trial. However, in considering one's opponent's choice on the last trial, it becomes apparent that if he or she is also rational, their last choice will also be to defect. Thus, behavior in the next-to-last trial has no possibility of influencing future behavior (on the last trial), and so forth. The only equilibrium in a game with a fixed endpoint is for both players to defect at every trial. In Rapoport and Guyer's (1966) terms, the equilibrium is deficient: A better joint outcome could be obtained by mutual cooperation on every trial.

When there is a nonzero probability of continuing the game after any trial, however, consideration of end-game play is less critical. Indeed, it becomes considerably less critical as the probability of continuing play increases. In the current study, a random number generated by the spin of a roulette wheel after each trial determined whether the game would continue. The probability that the game would continue was fixed and independent of the participants' choices.

EQUILIBRIUM INDICES

Roth and Murnighan (1978) derived two equilibrium indices that depend not only on the probability that the game will continue but also on the payoffs available to the players and the strategies that one's opponent is using. The indices are called equilibrium indices because they determine at what point for any particular game the cooperative choice is in equilibrium. (The individualistic, defect choice is always in equilibrium.) Unlike the indices for cooperation proposed by Rapoport and Chammah (1965), the equilibrium indices e_1 and e_2 may be interpreted as measuring the difficulty of achieving a cooperative equilibrium. The e_1 index determines the critical probability of continuing the game, below which no equilibria result in cooperation. The expected benefits from mutual cooperation do not outweigh the risks of one's opponent's possible defection when $p < e_1$. Only when $p > e_1$ does a cooperative equilibrium exist, and this can be achieved by the unforgiving strategy. The e_2 index has a similar interpretation in the context of equilibria involving the matching strategy. Specifically, the matching strategy for both players is in equilibrium if and only if the probability of continuing the game exceeds both e_1 and e_2. (For a more detailed discussion of these two indices, and of their relationship to equilibria of the game, see Roth and Murnighan, 1978). In the current study, then, if players behave rationally, they should cooperate less frequently if $p < e_1$ and they are faced with an unforgiving opponent or if $p < e_2$ or $p < e_1$ and they are faced with a matching opponent.

The e_2 index is identical to indices that have been proposed by several other theorists and researchers recently (Axelrod and Hamilton, 1981; Goehring and Kahan, 1976; Taylor, 1976) and appears to have been discovered independently by each of them. Goehring and Kahan reported data that supports its predictability in n-person prisoner's dilemma games that did not vary the probability of continuing. The Roth and Murnighan model, as far as we know, is the first to propose the e_1 index, and also is the first to demonstrate the relationship between the indices and the presence or absence of cooperative equilibria in a game. In addition to the model, Roth and Murnighan (1978) presented data from a single game that supported the predictions of both indices.

Table 3 displays the twelve games that were used in this study and the values taken by each of the indices for each of the games.

AN EXPECTED VALUE INDEX

Differentiating games on the basis of the presence or absence of a cooperative equilibrium ignores the magnitude of the gain or loss a cooperative strategy might produce. Formally, we wish to consider the expected gain (or loss) a player receives if he or she decides to play cooperatively in a game with probability p of continuing as compared to the expected payoff from playing noncooperatively (i.e., individualistically), against a given opponent.

The choice sequence that will occur in a given game is determined by the pair (s_1, s_2) of strategies for players 1 and 2, respectively, and by the random stopping time t at which the game terminates. Also associated with each strategy pair (s_1, s_2) is an expected payoff function $\pi(s_1, s_2) = (\pi_1(s_1, s_2), \pi_2(s_1, s_2))$, which gives the expected payoff $\pi_i(s_1, s_2)$ to player i (i = 1, 2) when player 1 plays strategy s_1 and player 2 plays strategy s_2.

Let u denote the unforgiving strategy, m the matching strategy, c the cooperative strategy that tells a player to cooperate in every period, the nc the noncooperative strategy that tells a player to defect in every period. Each player's payoff depends on the strategy pair (s_1, s_2) only to the extent that it determines the choice sequence; different strategy pairs yielding the same choice sequence yield the same expected payoff. Thus, for example, the strategy pairs (c, u) (c, m), (u, u), (m, m) and (c, c) all yield cooperative choices at every period, and so

$$\pi(c, u) = \pi_1(c, m) = \pi_1(u, u) = \pi_1(m, m) = \pi_1(c, c) = \quad [1]$$

$$a + \sum_{n=1}^{\infty} p^n a = \frac{a}{1-p}$$

Similarly, the strategy pairs (nc, u) and (nc, m) both yield the sequence in which player 1 makes the individualistic choice in every period, and player 2 makes the cooperative choice in period 0 only, and the individualistic choice thereafter. The expected payoff to player 1, which results from choosing the noncooperative strategy, is therefore

$$\pi_1(\text{nc, u}) = \pi_1(\text{nc, m}) = b + \sum_{n=1}^{\infty} p^n c = b + \left(\frac{cp}{1-p}\right) \qquad [2]$$

Thus, for a player whose opponent has adopted the strategy u or m, the expected gain (g) from cooperating (by choosing c, or u, or m) himself, as opposed to not cooperating (by choosing nc) is

$$g = \pi_1(\text{c, u}) - \pi_1(\text{nc, u}) = \frac{a}{1-p} - \left[b + \frac{cp}{1-p}\right] = \frac{1}{1-p}[a - b + p(b - c)] \qquad [3]$$

The g index is most easily compared to the two equilibrium indices by consulting Table 3. Cooperative equilibria for e_1 correspond to positive values for g. Thus, in the first game, when the probability of continuing is .50 or above, a cooperative equilibrium exists and $g \geqq 0$. When the probability of continuing is less than .50, no cooperative equilibrium exists and $g < 0$. The g index, then, discriminates among the game/probability conditions more finely than the equilibrium indices. It predicts, for instance, that there should be greater cooperation at each level of p in the fourth game in the table than in the eighth. From a psychological standpoint, the predictions of g are based on the simple assumption that the amount of cooperation observed will be correlated with the potential gains from cooperation.

The current study, then, addresses itself to several questions. The first is a test of the Roth and Murnighan equilibrium indices: Are they predictive of bargaining behavior in the set of twelve prisoner's dilemma games studied here? In addition, how do they compare to Rapoport and Chammah's and Thibaut and Kelley's indices in their ability to predict the frequency of cooperative choices of the players? Do players also cooperate most when the potential gains (g) from cooperation are

TABLE 3

The Games, e, r, and k Indices, and g Values for the Different Probabilities of Continuing

	Games				Indices										g Values for the Probability of Continuing		
	b	*a*	*c*	*d*	e_1	e_2	r_1	r_2	r_3	r_4	k_1	k_2	k_3	k_4	*.895*	*.50*	*.105*
(1)	40	39	38	0	.50	.026	.025	.975	.95	.025	−39	41	−37	117	8.0	0.0	−.89
(2)	40	39	35	0	.20	.026	.10	.975	.875	.025	−36	44	−34	114	35.0	3.0	−.56
(3)	40	39	30	0	.10	.026	.225	.975	.75	.025	−31	49	−29	109	80.0	8.0	0.0
(4)	40	39	20	0	.05	.026	.475	.975	.50	.025	−21	59	−19	99	170.0	18.0	1.11
(5)	40	30	29	0	.90	.33	.025	.75	.725	.25	−39	41	−19	99	−1.0	−9.0	−9.88
(6)	40	30	25	0	.67	.33	.125	.75	.625	.25	−35	45	−15	95	35.0	−5.0	−9.44
(7)	40	30	20	0	.50	.33	.25	.75	.50	.25	−30	50	−10	90	80.0	0.0	−8.88
(8)	40	30	15	0	.40	.33	.375	.75	.375	.25	−25	55	−5	85	125.0	5.0	−8.33
(9)	40	25	24	0	.93	.60	.025	.625	.60	.375	−39	51	−9	89	−6.0	−14.0	−14.89
(10)	40	25	20	0	.75	.60	.125	.625	.50	.375	−35	55	−5	85	30.0	−10.0	−14.44
(11)	40	25	15	0	.50	.60	.375	.625	.375	.375	−30	60	0	80	120.0	0.0	−13.33
(12)	40	25	1	0	.30	.60	.60	.625	.025	.375	−16	74	14	66	201.0	9.0	−12.33

greatest? The effects of the independent variables can also be observed (1) as the probability of continuing play increases, (2) when one's opponent uses the unforgiving strategy rather than the matching strategy, or (3) when the players are males or females. Finally, the effects of expecting either cooperative choices by one's opponent or a higher probability of continuing play can be correlated with one's choices to see if the relations are positive, as expected (Wyer, 1969). The effects of the independent variables on these estimates can also be ascertained.

METHODS

Subjects. Participants were 252 male and female undergraduates who completed part of a course requirement by participating. The students, all enrolled at the University of Illinois at Urbana—Champaign, chose this study from several available at the time.

Design. Participants were randomly assigned to one of the twelve prisoner's dilemma games shown in Table 2 and to one of two opponent's strategies. Half the players faced an opponent who used the matching strategy. The remaining players faced an opponent who used the unforgiving strategy.

All players played a single game three times with three "different" opponents, with the probability that the game would continue (p) varying from one game to the next. Previous research (Roth and Murnighan, 1978) indicated that the order of play had no effect on the player's choices. Thus, the players played the three games in the same order. For the first game $p = .895$; for the second, $p = .50$; and for the third, $p = .105$. Thus, the design was mixed: games and opponent strategies were between-subject factors; probability of continuation was a within-subject factor. In addition, the numbers of males and females in each of the game-opponent strategy conditions were approximately equal. Due to scheduling difficulties, it was not possible to keep cell sizes exactly equivalent.

The primary dependent variable in the analyses was the proportion of cooperative choices in each of the games. Secondary analyses were also conducted on the players' first choices in each game. Because of the probabilistic termination of the games, an unequal number of trials resulted across individuals and games. The proportions and first choices, however, can be compared across various sessions. (The proportions were given priority in the analyses because of their greater stability.) All of the analysis of variance effects that included the probability of continued play as a factor were subjected to the conservative test proposed by Box (1954) and discussed by Winer (1962). This approach is a statistical attempt to control for the possible nonindependence of effects due to the repeated nature of the probability factor. The degrees of freedom for p are simply reduced from two to one in all significance tests.

Other variables of interest, measured prior to each choice, were the subject's perceptions of the probability that play would continue for at least two additional trials [p(2T)] and the probability that one's opponent would choose cooperatively [p(C)]. Finally, an eleven-item questionnaire designed to tap individualistic versus collectivistic motivations was also administered.

Procedure. Subjects were run in groups of five or six. Instructions were read to the entire group. Participants were encouraged to try to maximize their own payoffs, without concern for their opponents' outcomes. The payoff matrix to be used during the sessions was discussed, and the choices that the players could make at each point in the games were emphasized, as was the determination of the outcomes of each pair of subject's choices. Players were also instructed that they would have practice sessions prior to each game, and that they would be asked to indicate how likely they felt it would be (in terms of probabilities) that the other player would choose A, the cooperative choice, and how likely it would be that the game would continue for at least two additional trials. Use of the roulette wheel to determine whether a game would continue was also demonstrated and discussed.

After answering all questions about the procedure, subjects were taken to another room and seated behind opaque partitions that isolated each subject from view of all other participants. They were told that they would be playing a different individual in each of the three sessions, but that the person's identity would not be revealed. Actually all of the subjects played against the experimenter, who implemented either the matching or unforgiving strategy.

Subjects first practiced making the cooperative and defect choices by sliding cards marked A and B under the partitions to the experimenter. Prior to each choice, subjects recorded their perceptions concerning the likelihood of two additional trials and the probability of their opponent choosing A. They also recorded their own choice and its outcomes, which was fed back to the players in written format. Thus, there was no verbal discussion during the course of the experiment. Following several practice choices, subjects were told that they would be switching partners (opponents), and that the probability that they would continue play with the new partner would be determined by a spin of the roulette wheel after each pair of choices. They were given approximately five choices during the practice phase prior to playing the game with a probability of continuing of .895. They were told that this probability was determined by the fact that there were 38 slots in the roulette wheel (numbers 1 through 36 plus 0 and 00) and that if the ball fell in slots 0, 00, 1, or 2, the game would be terminated. If the ball fell in slots 3 through 36, the game would continue. The roulette wheel was spun after each play of the game, for both practice and "money" trials (see below).

These initial practice sessions were run to familiarize the subjects with the payoff matrix, the choices they could make, the outcomes they could receive, the effect of the roulette wheel, and the questions they were asked to answer prior to making their choice. After the practice trials, subjects were told that the money trials would begin. The money trials were conducted in the same way as the immediately preceding practice trials with the roulette wheel, but the subjects' outcomes now determined whether they would be the winners of one of three $40 first prizes or three $20 second prizes. The players were told that in each of the three money games, the student who obtained the best average score per trial would receive $40 after completion of the study. The second best average score would receive $20. Separate $40 and $20 payoffs were awarded for each of the three games.[2]

The second and third sessions of the experiment presented the subject with an ostensibly new partner, the same payoff matrix, and different probabilities that the game would continue. For the second session $p = .50$: the game was terminated if the roulette ball fell in slots 0, 00, or 1 through 17. For the third session $p = .105$: the game was terminated if the ball fell in slots 0, 00, or 1 through 32. As with the first session, practice trials were run with the roulette wheel prior to the money trials.

Following completion of the three games, subjects were asked to fill out a short questionnaire designed by Breer and Locke (1965) to tap

2. This procedure was adopted because of financial constraints. In consequence, the players' behavior may be less risk-averse than if their payoffs were linear in their point score. For a discussion of how experimental designs can be implemented which control for the risk-aversion of the players, see Roth and Murnighan (1982).

their individualistic and collectivistic motivations. A question concerning their prior familiarity with prisoner's dilemma games was also included. Finally, subjects were told the purposes of the study, and the deception concerning their opponents was revealed. Only a few subjects expressed any suspicion that they were not actually playing one of the other people at the experiment, and none objected to the deception. Questions were answered, and subjects were given written feedback including the experimenters' names as sources for additional information. Upon completion of the study, the results were tabulated to determine the prize winners. The prizes were awarded to the players who maximized their point totals, controlling for payoff matrices for a particular game.

RESULTS

The results will be presented in several sections. After reviewing the effectiveness of the manipulations and the consistency of the players' choices, the analysis of the effects of the independent variables will be presented. These will be followed by analyses of the predictability of the different models and correlational data.

MANIPULATION CHECKS

Subjects were verbally quizzed prior to the start of the experiment to determine whether they understood how the payoffs would be determined. Everyone seemed to understand the matrices before play began.

The opponent's strategy manipulation can be checked to determine whether different strategies led to different estimates of the probability that one's opponent will choose cooperatively (C), especially after the subject chose individualistically. Indeed, although every opponent chose C on the first choice, subjects predicted a much lower likelihood of a C response (mean P(C) = .49). After subjects made their first D response, their estimates of their opponent's cooperative choices went down (mean P(C) = .27), but did not drop to zero. In addition, there were no significant differences for subjects facing the matching versus the unforgiving strategy. This variable, then, appears to have had little effect on the subjects' perceptions. Not surprisingly, it also had little effect on the other dependent variables.

Subjects' estimates of the probability that there would be at least two additional trials showed strong effects for the independent manipulation of the probability that the game would continue ($F(1,408) = 672.62$, $p < .001$). When p equalled .895, subjects mean estimates were .844; when p = .50, subjects' estimates were .492; when p = .105, subjects' estimates were .183. Clearly, this manipulation was very effective, even though the subjects overestimated the correct probabilities (.80, .25, and .01).

A final check on the sample of participants concerned their familiarity with prisoner's dilemma games. Although some had experienced material on the game in classes, none had actually played a prisoner's dilemma game. Correlating familiarity with all other variables in the study led to no relationships that approached significance. Also, none of the other questionnaire items showed any significant relationship with any of the other variables.

CONSISTENCY OVER TRIALS

Because of the varying number of trials, the consistency of the players' responses as the trials progressed must be addressed. Table 4 summarizes players' choices in the early and later trials. Some response variation can be observed in several of the games. While players responses were consistent, for instance, in game 7, they were not so consistent in game 8. Thus, the use of the proportion of a player's cooperative choices as the primary dependent variable in subsequent analyses is indicated. Table 4 also shows that players in the .50 and .105 conditions cooperated less as the trials progressed; this is not apparent in the .895 game.

EFFECTS OF THE INDEPENDENT VARIABLES

In a 12 (games) × 2 (opponent strategy) × 2 (sex) × 3 (probability of continuing) analysis of variance, with games, opponent strategy, and sex between variables and probability of continuing a repeated, within-variable, the percentage of cooperative choices showed the following significant effects: games ($F(11,204) = 2.11$, $p < .03$); probability ($F(1,408) = 23.07$, $p < .01$); and games by probability ($F(11,408) = 2.20$, $p < .05$). The effects for players' first choices were very similar, though somewhat reduced in magnitude. The means for the main effects and interaction are shown in Table 5.

TABLE 4

The Mean Number of Trials and Percentage Cooperation (%C) on Early Trials in the Game-Probability Conditions

Probability of Continuing =					.895					.50					.105		
	Game				*Mean No.*	*%C*	*%C*	*%C*	*%C*	*Mean No.*	*%C*	*%C*	*%C*	*%C*	*Mean No.*	*%C*	*%C*
	b	*a*	*c*	*d*	*of Trials*	$t = 1$	$t = 2$	$t = 2$	$t > 4$	*of Trials*	$t = 1$	$t = 2$	$t = 3$	$t = 4$	*of Trials*	$t = 1$	$t = 2$
(1)	40	39	38	0	10.20	26.1	34.8	30.4	30.0	1.81	47.8	46.2	–	–	1.14	39.1	33.3
(2)	40	39	35	0	9.89	42.1	21.1	26.3	11.1	2.74	31.6	26.3	7.1	–	1.26	21.1	20.0
(3)	40	39	30	0	10.57	34.8	52.2	43.5	44.4	1.57	52.2	38.5	–	–	1.00	34.8	–
(4)	40	39	20	0	10.20	25.0	35.0	50.0	36.4	2.20	55.0	33.3	66.7	–	1.00	30.0	–
(5)	40	30	29	0	10.04	33.3	20.8	33.3	34.7	2.21	33.3	15.8	0.0	20.0	1.00	16.7	–
(6)	40	30	25	0	10.25	30.0	50.0	45.0	33.3	2.55	30.0	35.0	27.3	–	1.00	10.0	–
(7)	40	30	20	0	10.35	25.8	25.8	25.8	21.9	2.35	38.7	38.7	18.2	–	1.00	19.4	–
(8)	40	30	15	0	9.67	16.7	11.1	38.9	37.0	2.00	50.0	27.8	40.0	–	1.00	0.0	–
(9)	40	25	24	0	9.10	10.0	20.0	15.0	22.7	2.55	5.0	0.0	13.3	0.0	1.25	15.0	0.0
(10)	40	25	20	0	8.00	31.6	26.3	28.6	20.0	2.00	15.8	21.1	–	–	1.53	5.3	0.0
(11)	40	25	10	0	9.00	29.4	41.2	58.8	58.8	2.00	41.2	30.8	25.0	–	1.00	5.9	–
(12)	40	25	1	0	10.22	44.4	55.6	61.1	47.5	2.28	50.0	44.4	60.0	–	1.00	16.7	–

TABLE 5

The Means and Simple Effects for the Game X Probability of Continuing Conditions for the Percentage of Cooperative Choices Over All Trials

	Games				*Probability of Continuing*				*Simple Effects*		
	b	*a*	*c*	*d*	*.895*	*.50*	*.105*	*M*	*F*	*df*	*p* $<$
(1)	40	39	38	0	32^{++}	46^{++}	41^{--}	40	1.52	1,44	ns
(2)	40	39	35	0	24^{++}	23^{++}	21^{--}	23	<1	1,36	ns
(3)	40	39	30	0	43^{++}	52^{++}	35^{++}	43	2.27	1,44	ns
(4)	40	39	20	0	35^{++}	48^{++}	30^{++}	38	2.34	1,38	ns
(5)	40	30	29	0	31^{--}	24^{--}	21^{--}	26	<1	1,46	ns
(6)	40	30	25	0	27^{++}	31^{--}	10^{--}	26	5.12	1,38	.05
(7)	40	30	20	0	24^{++}	38^{++}	16^{--}	26	4.78	1,60	.05
(8)	40	30	15	0	28^{++}	39^{++}	0^{--}	22	9.48	1,34	.01
(9)	40	25	24	0	16^{--}	5^{--}	12^{--}	11	<1	1,38	ns
(10)	40	25	20	0	28^{++}	18^{--}	3^{--}	16	3.97	1,36	.06
(11)	40	25	10	0	46^{++}	38^{+-}	6^{--}	30	10.20	1,32	.01
(12)	40	25	1	0	52^{++}	49^{+-}	17^{--}	39	9.00	1,34	.01
			M		33	34	18	28			

NOTE: Superscripts indicate when a particular probability/game combination yields a cooperative equilibrium: ++ indicates that a cooperative equilibrium exists for e_1 and e_2; +– for e_1 not e_2; and –– indicates no equilbrium for either e_1 or e_2.

Post hoc tests using the Newman-Keuls procedure (α = .05) showed that the players made significantly fewer cooperative choices (18%) when p = .105. The effects for games and the interactions between games and probability are not as strong as the main effect for probability, but the interaction suggests that the pattern of results changes as a changes: when a = 39, there is little or no difference among game/probability combinations; when a = 30, cooperation decreased as c decreased when p = .105; and when a = 25 cooperation decreased as c increased or as p decreased.

Sex also had significant effects on players' perceptions: Females felt that two trials were more likely than males (mean $X_{females}$ = .54; mean X_{males} = .47; $F(1,204) = 11.02$, $p < .001$) and were more likely than males

TABLE 6
Summary of Main Effects and Interactions of the Indices on the Percentage of Cooperative Responses

Index	F_{Index}	*df*	$p <$	$F_{Index\ X\ Prob}$	*df*	$p <$
r_1	1.76	3,236	ns	4.86	3,472	.01
r_2	3.47	2,204	.04	4.00	2,408	.03
r_3	2.28	2,240	.11	6.26	2,480	.01
r_4	3.47	2,204	.04	4.00	2,408	.03
e_1	5.66	2,240	.005	1.98	2,480	ns
e_2	3.47	2,204	.04	4.00	2,408	.03
g	7.77	4,736	.001	–		
k_1	1.86	3,236	ns	4.66	3,472	.01
k_2	3.36	2,240	.04	3.77	2,480	.04
k_3	2.50	2,240	.10	6.24	2,480	.01
k_4	2.50	2,240	.10	6.24	2,480	.01

NOTE: The results for r_2, r_4, and e_2 are identical: they are perfectly correlated indices in this study. The same is true for k_3 and k_4.

to think that their opponent would choose C (mean $X_{females} = .53$; mean $X_{males} = .45$; $F(1,204) = 6.19$, $p < .02$).

PREDICTIONS OF THE INDICES

Each of the indices were used as independent variables in combination with opponent strategy, sex, and, except for g, probability of continuing, in separate analyses of variance with the three dependent variables. Each index was partitioned into a smaller set of levels to increase the cell sizes within each analysis; the values chosen for the different levels limited the index's range as much as possible within each level. Index, opponent strategy, and sex were between variables; probability of continuing was the only repeated measure.

Table 6 summarizes the analysis of variance results for the main effects and interactions of the indices with the probability of continuing for the percentage of cooperative responses. (Again, analyses of players' first choices closely approximated these results.) All of the indices except r_1 and k_1 showed significant or marginally significant main effects, and almost all the indices interacted significantly with the probability of continued interaction. The data support some models more than others: As r_1, r_2, k_1, k_2, k_4, and g increase or as r_4, e_1, e_2, and k_3 decrease, the percentage of cooperative choices increase. For moderate values of r_3, however, cooperation was lower than for either high or low values. The significant interactions with the probability of continuing support the predictions of the equilibrium indices, particularly e_2. Indeed, the condition where cooperative equilibria existed (see Table 5) yielded rates of cooperation that were higher in almost every case than the conditions where cooperative equilibria did not exist. In other words, when the probability of continuing exceeded the equilibrium index, the mean percentage of cooperative choices was greater than when the probability of continuing did not exceed the index.

CORRELATIONAL DATA

Correlational analyses can also be utilized to evaluate the predictive accuracy of the indices. Each index was correlated (using Spearman's rho) with both players' first choices and the percentage of their cooperative choices across games for each of the three probabilities of continuing (g was correlated with these two variables across all three probabilities). The results are displayed in Table 7. In general, the indices did poorly for the .895 conditions. The r_1, r_3, and k indices all do poorly in one of the other two conditions as well. This suggests some (not overwhelming) superiority for r_2, r_4, e_1, and e_2. (Note that the results for e_2 are identical to those for r_2 and r_4.) The results for g surpassed those of the other indices: For first choices, rho = .18; for the percentage of cooperative choices, rho = .30. Both figures exceed the average correlations of all of the other indices.

Finally, as further predictors of the rate of cooperation, subjects' estimates of the likelihood that their opponent would choose cooperatively and the likelihood of two additional trials were correlated with the cooperation rate. The probability estimates of two additional trials were not significantly related to estimates of the likelihood of the opponent's cooperation or the subjects' own cooperation rates. Subjects' perceptions that their opponent would choose C, however, correlated significantly with their own cooperation, supporting previous findings. In addition, as the probability of continued play decreased, the correlations also decreased. When p = .895, r = .41; p = .50, r = .37; p =

TABLE 7
Spearman Rank Order Correlations (ρ) Between First Choices and Percentage Cooperation and Each of the Indices

Probability of Continuing =	*.895*		*.50*		*.105*	
Index	*First Choice*	*Percentage Coop*	*First Choice*	*Percentage Coop*	*First Choice*	*Percentage Coop*
r_1	.03	.15	.17	.19	–.06	–.09
r_2	.03	–.03	.17	.14	.22	.24
r_3	.02	–.10	–.02	–.04	.19	.21
r_4	–.03	.03	–.17	–.14	–.22	–.24
e_1	–.07	–.11	–.24	–.23	–.13	–.12
e_2	–.03	.03	–.17	–.14	–.22	–.24
k_1	..03	.15	.17	.19	–.06	–.09
k_2	.03	.15	.17	.19	–.06	–.09
k_3	–.03	.07	–.07	–.04	–.23	–.25
k_4	.03	–.07	.07	.04	.23	.25

NOTE: n = 252 for each correlation. Values greater than .10 or less than –.10 are significant at $\alpha = .05$; values greater than .15 or less than –.15 are significant at $\alpha = .01$.

.105, r = .21. Thus, as continued interaction becomes less likely, perceptions of an opponent's choices influence an individual's choice less and less.

DISCUSSION

The results emphasize the importance of considering the probability of continued play in the study of prisoner's dilemma games. Before discussing this variable, however, some discussion of the other variables of interest in this study is warranted.

Both the strategy of one's opponent and the sex of the players had little effect in the present study. Although the procedures were designed to make the participants aware of their opponent's choices and the pattern of those choices, subjects did not seem particularly aware of the differences between the matching and unforgiving strategies. This may have been due to the relatively small number of choices made prior to each of the money trials and the fact that these choices were "practice." Although subjects recorded their own choices and the outcomes of each of the practice trials, they may not have converted this information into a profile of their opponent. In addition, the players experienced a fairly small number of trials overall, which could have made it more difficult for them to detect or respond to a particular opponent strategy.

As with much of the research on prisoner's dilemma games, the differences found between the male and female participants were not particularly illuminating. Females expected others to be more cooperative and were more cooperative than males, but not significantly so. As has been mentioned before, systematic research into the dynamics of prisoner's dilemma situations is needed to determine what variables might interact with the sex of the players to determine their choices (e.g., Hartman, 1980). The current study included sex as a variable more for control than for diagnostic reasons.

The fact that the probability of continued interaction was both accurately perceived by the players and also very strong in affecting their behavior supports an underlying assumption of the Roth and Murnighan (1978) model. The results showing p to interact with games is also inherent in the predictions of the equilibrium models. The other models performed quite well across the different levels of p; they were not derived, however, for games with probabilistic termination points. Their usefulness depends, then, on the salience of the expectation of continued play in games similar to the prisoner's dilemma. If games do not have fixed termination points, or if individuals expect their interactions to end in ways that can be modeled by a fixed probability, then this approach will better reflect the strategic considerations of players in mixed motive games. The study of probabilistic termination, in our opinion, moves the current research toward a more realistic analog to "real world" games. It also provides the basis for an empirical test of the equilibrium models.

How accurate were the equilibrium indices in predicting behavior? The e_2 index made identical predictions to Rapoport and Chammah's second index and to Harris's r_4 index. The significant effects for probability clearly favor e_2. While it appears, then, to make predictions that are similar to those of r_2 and r_4, it improves on them by differentiating among conditions with different termination probabilities. The e_1 index was also supported. As with e_2, cooperation was higher in every case where a cooperative equilibrium existed.

What about g? It makes intuitive sense to think that, when cooperation gives an individual a larger expected payoff, he or she will tend to be more cooperative than when it offers a smaller expected payoff. The data do suggest that variations in g had an impact on the outcomes, but not in the anticipated pattern. The correlational data strongly support g: It correlated more strongly with cooperative choices than any of the other models. The analysis of variance results, however, are less clear. Instead of observing a continuous increase in cooperation as g increased, a clear step function seems to have emerged. When g is close to or larger than zero, cooperation rises to its behavioral limit (given the conditions of this experiment). This pattern is much more in line with the underlying assumptions of the two equilibrium models. Thus the differentiation provided by g may be more than is necessary. Or the potential payoffs may have influenced the results more than originally anticipated, for the payoff function in this study also was a step function. It does suggest that, in the future, more substantial monetary gains might be associated with the outcomes to ensure that participant's perceptions of their expected payoffs more clearly fit the expected payoffs used by g to predict behavior.

CONCLUSION

This study presented a large set of data that investigated a new variable, the probability of continued play, in the study of prisoner's dilemma games. In combination with different game payoffs, this variable significantly affected the rates of cooperation observed. In particular, as the probability of continued play increased, for particular games, cooperation also increased. The pattern of behavior observed was modeled quite well by two equilibrium indices (Roth and Murnighan, 1978). The article also introduces a new model, g, which is based on the expected payoffs that can be obtained (or lost) from cooperating. This index was more highly correlated with cooperative choices than any other index. Thus, both the original equilibrium model and its indices, e_1 and e_2, and the new model, g, have been supported more than the previous models and their indices. A final conclusion from this article is that, in future research on mixed-motive games (of which the prisoner's dilemma is the most often studied), the expectation of continued play might be an important determinant of cooperative behavior, and that such an expectation may interact with the payoffs available.

REFERENCES

AXELROD, R. (1980a) "Effective choice in the iterated prisoner's dilemma." J. of Conflict Resolution 24: 3-25.

——— (1980b) "More effective choice in the prisoner's dilemma." J. of Conflict Resolution 24: 379-403.

———and W. P. HAMILTON (1981) "The evolution of cooperation." Science 211: 1390-1396.

BOX, G.E.P. (1954) "Some theorems on quadratic forms applied in the study of analysis of variance problems." Annals of Mathematical Studies 25: 290-302, 484-498.

BREER, P. E. and E. A. LOCKE, (1965) Task Experience as a Source of Attitudes. Homewood, IL: Dorsey.

GOEHRING, D. J. and J. P. KAHAN, (1976) "The uniform n-person prisoner's dilemma game." J. of Conflict Resolution 20: 111-128.

HARRIS, R. J. (1969) "A geometrical classification for 2x2 interval-symmetric games." Behaviorial Sci. 14: 138-146.

HARTMAN, E. A. (1980) "Motivational bases of sex differences in choice behavior." J. of Conflict Resolution 24: 455-475.

JONES, B., M. STEELE, J. GAHAGAN, and J. TEDESCHI (1968) "Matrix values and cooperative behavior in the prisoner's dilemma game." J. of Personality and Social Psychology 8: 148-153.

OSKAMP, S. (1971) "Effects of programmed strategies on cooperation in the prisoner's dilemma and other mixed-motive games." J. of Conflict Resolution 15: 274-283.

RAPOPORT, A. and A. M. CHAMMAH, (1965) Prisoner's Dilemma. Ann Arbor: Univ. of Michigan Press.

RAPOPORT, A and M. GUYER (1966) "A taxonomy of 2 × 2 games." General Systems 11: 203-214.

ROTH, A. E. and J. K. MURNIGHAN (1982) "The role of information in bargaining: an experimental study." Econometrica 50: 1123-1142.

——— (1978) "Equilibrium behavior and repeated play of the prisoner's dilemma." J. of Mathematical Psychology 17: 189-198.

STEELE, M. W. and J. T. TEDESCHI (1967) "Matrix indices and strategy choices in mixed-motive games." J. of Conflict Resolution 11: 198-205.

TAYLOR, M. (1976) Anarchy and Cooperation. London: John Wiley.

THIBAUT, J. W. and H. H. KELLEY (1959) The Social Psychology of Groups. New York: John Wiley.

WINER, B. J. (1962) Statistical Principles in Experimental Design. New York: McGraw-Hill.

WYER, R. S. (1969) "Prediction of behavior in two-person games." J. of Personality and Social Psychology 13: 222-238.