GENERAL SYSTEMS
Volume XX, 1975

BAYESIAN MODELS FOR ITERATED PRISONER'S DILEMMA GAMES*

Bernard Grofman and Jonathan Pool

The notion of "rational choice" does not appear clear cut for two-actor decision problems other than those involving pure risk (where one of the actors is Nature) or for games other than constant-sum, save when there exists a dominant strategy for at least one of the players. We believe that choice of a dominant strategy is always rational.[1] Several scholars, most notably Anatol Rapoport (1960, 1964, 1966; Rapoport and Chammah 1965), have argued that choice of a dominant strategy need not be "rational" and have buttressed this argument by referring to the Prisoner's Dilemma game. This game has been very widely discussed, and a voluminous literature has grown up around it. *All* recent articles seeking to refer game theory to ethics discuss it (e.g., Wolff 1963, Thompson 1964, Hopkins 1965, Held 1966, Runciman and Sen 1965, Cunningham 1966, Schelling, 1969). Hyperbolic claims have been made as to its importance for social ethics and for the theory of rational choice. Anatol Rapoport and Albert Chammah have likened the "paradox" of the Prisoner's Dilemma to the "paradoxes of relativity and of quantum mechanics," forming the "likely basis of a far-reaching philosophic reconstruction" (1965, p. 12).

Unfortunately, much of the literature on the game has been marred by failure to distinguish between the properties of the game matrix itself and the connotations attached to the game matrix by virtue of association with the morality tale of the Prisoner's Dilemma customarily associated with the game. Thus the "game" has been asserted to have demonstrated that selfishness (concern only for one's own payoffs) and lack of trust are socially undesirable in that they lead (or at least can lead) to outcomes which neither player wants and which can be avoided if only the players would trust each other or if only each could incorporate into their utility function some concern for the payoffs to the other player, assertions which we regard as almost entirely erroneous (See Grofman 1975).

A Prisoner's Dilemma game matrix is shown in Figure 1.[2] It is asserted that a paradox arises in consequence of the fact that if both players choose their dominating strategies, the outcome (d, d) results, which is inferior for

		Column 1	Column 2	
Row	1	(a, a) E_{11}	(b, c) E_{12}	I $c>a>d>b$
	2	(c, b) E_{21}	(d, d) E_{22}	II $2a>b+c>2d$

Fig. 1. Symmetric Prisoner's Dilemma Game[a,b,c]

a-Alternative 1 is normally labeled "cooperation" and alternative 2 is normally labeled "defection."

b-Our defining characteristics for the P.D. game are those given in Rapoport and Orwant (1962) plus the usual stipulation of symmetry. As Oskamp (1970) has pointed out, many games labeled as P.D. in the experimental gaming literature do not, in fact, satisfy conditions I and II.

c-This figure employs standard notation; i.e., "Row" and "Column" are the two players, E_{ij} is the outcome when Row chooses alternative i and Column chooses alternative j; (i, j) shown over an outcome indicates the pay-offs to Row and Column, respectively, associated with that outcome.

both players to the outcome which would have been obtained had each behaved "irrationally" (i.e., had chosen strategy 1).

In this paper we shall be interested in showing some results of models of choice in infinitely iterated Prisoner's Dilemma games which are inspired by notions of maximizing long-run expected payoff based on subjective estimates of the response probabilities of the other player. First, however, we shall briefly discuss various "solutions" to the P.D. game proposed in the literature. As Anatol Rapoport (1970, p. 177) notes, "different people have different ideas of what constitutes a 'solution'." By a "solution" we shall here mean simply an attempt to specify a "reasonable" strategy for the game.

We note first that it is necessary to distinguish between purported solutions to the P.D. game when played non-cooperatively and those to the game played cooperatively and to relate the latter to a wide range of communication and agreement options; also to distinguish among proposed solutions to one-shot, finitely iterated, and to infinitely iterated P.D. games; further, to dis-

*This paper draws heavily on Bernard Grofman's "Rationality and Choice in the Context of Two-Person Non-Cooperative Games" (unpublished Ph.D dissertation, University of Chicago, 1972). An earlier verson of the first part of this paper was presented by Prof. Grofman at the National Science Foundation Seminar on "The Mathematical Theory of Collective Decisions," Hilton Head Island, South Carolina, August 8-14, 1971.

1. A defense of this view is found in Grofman (1975).

2. Accompanying this matrix there is a vignette to account for the properties of the payoff outcomes: two prisoners, partners in crime, are locked in separate cells unable to communicate with each other, after having been arrested separately and unexpectedly. If only one prisoner turns state's evidence, he will get off scot free (payoff c), while the other prisoner will receive a very harsh sentence (payoff b). If both confess, then each gets a long sentence (payoff d) but not so long as that of the prisoner who didn't turn state's evidence while his partner did. If neither confess, they both get off with a mild sentence for an aggravated case of spitting on the sidewalk (payoff a), the only offense the cops can pin on them in the absence of a confession by either or both. In the absence of communication between the prisoners and of binding (near costless) contracts they could enter into, no matter what his partner does, it is always rational (utility maximizing) for the other prisoner to turn state's evidence and confess. (It is assumed that all the values of each are reflected in their payoffs which are assumed to satisfy criteria I and II of Figure 1.) This is true even if one (or both) partners have complete trust in the other's not confessing.

tinguish among P.D. games with various stop rules, discount rates, payoff structures, survival considerations, etc. We agree with Martin Shubik's assessment of the P.D. literature (Shubik 1970, p. 191), that often a class of games has been taken for a single game and too simple a construct is used to explain too much. We shall confine our discussion to solutions that have been proposed for strictly non-cooperative[3] versions of the P.D. game, where the entire game matrix is known to both players. We shall begin with proposed "solutions" to the one-shot Prisoner's Dilemma.

"Solutions" to One-Shot Prisoner's Dilemma Games

First, consider the *maximin* solution. Whereas maximin considerations may not be especially relevant in a non-constant sum context, in one-shot P.D. they are reinforced by considerations of dominance, the so-called "sure thing" principle. Following Luce and Raiffa (1967, p. 96 with change in notation), "We do not believe that there is anything irrational or perverse about the choice of strategy 2 and we must admit that if we were actually in this position we would make these choices." However, while both the maximin choice and the choice of the dominant strategy lead to the same outcome in one-shot P.D., it is important to distinguish the two solution rationales, as Rapoport and others sometimes do not. The importance of the distinction is that, in iterated games, the iterated maximin choice need not be the dominant one.

A third "solution" to the one-shot P.D. game is the proposal that each player "ought" to incorporate the other player's payoffs into his own utility calculations, that is, should seek to maximize joint payoff (and *presumably* joint utility) rather than his own. In this connection, Rapoport (1970) distinguishes between the *individually* rational solution and the *collectively* rational solution to the one-shot P.D. game. It is not clear why it is rational (as opposed to moral) for an actor to weigh the other's payoffs equally with his own. If he weighs them less than his own, we have shown elsewhere (Grofman 1975) that the P.D. situation may persist even in the transformed utility matrix. Of course, it may be quite rational (from the perspective of the authorities) to seek to inculcate consideration for others among members of the society. But such other-regarding behavior need not necessarily be considered "rational." It should be pointed out in fairness to Rapoport, that, in his view in the context of the P.D. game, "rationality" has two irreconcilable meanings.

The fourth "solution" is also based on moral grounds. It is based on Kant's generalization agrument, popularly phrased in the form

> ". . . But suppose everyvody did the same!", the generalization argument calls for appealing to the consequences of everybody's doing X (where X stands for an identifiable kind of act). For example, if I tell you that I don't think I'll bother voting in the next election, you might use the generalization argument to try to show me that I would be wrong in not voting. You say to me: "But suppose everybody did the same, suppose everybody decided not to vote; then the result would be disastrous. . . ." An adherent of the generalization argument must on moral grounds choose the course of action, which chosen by everybody, would have the best consequences (Cunningham 1966, p. 15).

There are two principal difficulties with this line of argument. First, to choose on the basis of the argument, "What if everybody did the same?" is nonsensical unless there is some direct causal nexus between your actions and the actions of others. In the one-shot non-cooperative P.D. game in which the players choose simultaneously, no such causal nexus exists. Second, while it may be that following self-imposed limitations is the *moral* thing to do, unless such concerns so motivate the actor as to transform a P.D. payoff matrix into one that is no longer a P.D., they are irrelevant to "rational" choice. Thus neither the "collective rationality" solution nor the generalization argument really "solve" the P.D. game; they only avoid it. (Cf. Cunningham, 1966.)

Let us now turn to a fifth proposed solution to the one-shot P.D. game which was once hailed by Anatol Rapoport as *the* solution to the one-shot P.D. game (Rapoport 1967), namely Nigel Howard's (1966a, 1966b, 1970) theory of metagames and meta-game rationality.[4] In Howard's model, the collectively rational strategy also becomes individually rational, since it becomes an equilibrium point in the metagame, which it was not in the original game. In Rapoport's words (1970, p. 173), "The [metagame] device is a generalization: a metastrategy is to strategy as a strategy is to a move in a sequential-move game. By 'ascending' into the space of metastrategies, one obtains a new perspective of the game."

To see how Howard's metastrategies operate, instead of considering only two alternatives for one of the players (say, column), let us imagine four: (11) play 1 regardless; (12) play 1 (2) if you think Row will play 1 (2); (21) play 1 (2) if you think Row will play 2 (1); (22) play 2 regardless. The four strategies of column are divided into two types: *conditional* (upon expectations of the other player's behavior) and *unconditional*. This expanded strategy space for column may be represented in a new meta-game matrix as in Figure 2. Since 22 dominates 11 and 12 dominates 21, we may eliminate strategies 21 and 11. In the reduced matrix it is apparent that the dominant strategies are 12 and 2 and the equilibrium point is the familiar (d, d). But, as Howard puts it (1966a, p. 169),

		Column			
		11	12	21	22
Row	1	(a, a)	(b, c)	(a, a)	(b, c)
	2	(c, b)	(d, d)	(d, d)	(c, b)

Fig. 2. Game Matrix for First Order Metagame Strategies.

3. By a non-cooperative game we shall here simply mean one in which there is no communication between the players other than (in an iterated game) that conveyed by the reports to each of the other's previous choices.

4. Rapoport has more recently retreated from use of the definite article after having been criticized for it by Martin Shubik (1970, p. 190). Rapoport states that "since Shubik takes exception to my view that Howard's model is the solution of the paradox, I feel that some clarification is called for. My use of the definite article was perhaps unjustified; different people have different ideas of what constitutes a 'solution' of any situation that calls for one. I do think that Howard's metagame theory attacks the core of the paradox as other proposed solutions do not" (Rapoport 1970, p. 177).

"This is not the end of the matter, indeed there is no end to it." The next step is to generate an expanded set of strategies as follows. Consider Row contemplating the matrix of Figure 2. In like manner, as above, we may generate a total of (4^2) alternative conditional and unconditional strategies for Row which can be denoted 1111, 1112, 1121, 1211, 2111, 1122, 1212, . . . 2222; which stand for

> 1111–play 1 regardless
> 1112–play 1 unless you think column will select 2 regardless of what he thinks your choice is
> .
> .
> .
> 2222–play 2 regardless

A 16 × 4 matrix for the 2nd order metagame results from the enlargement. There are now three equilibrium points, two of which are the jointly optimal (a, a), and one of which is the familiar (d, d). Thus, according to Howard, if Row makes the only choice which is always rational and column reacts rationally, the results will be the cooperative outcome (a, a). "The dilemma is solved" (Howard 1966a, p. 180). Moreover, Howard proves (1966b) that it is unnecessary to go any further in constructing $2^{n+1} \times 2^n$ metagames for metalevels $n > 2$. "[I]n a two person game all equilibria are found on reaching the second 'meta-level' " (op. cit., p. 169).

Let us, as Howard puts it, "Note how the (a, a) solution comes about."

> It is, of course, an irrational solution for each player given the other's choice. But it is a rational solution for each player given the other's meta-choice. It comes about through [Column] asking himself whether he should be rational, or whether some other policy might be rational in view of the fact that [Row] can ask himself the same question. (Howard 1966a, p. 180)

Howard regards his theory as a "static equilibrium theory." That is, the equilibria may be regarded as the possible results of a dynamic process. The latter is not formally described in the theory but may be informally inferred. Howard goes on to suggest a number of elements which might go to make up a dynamic process including iterations, bargaining, public commitments, threats, etc. *None of these, however, is relevant to the one-shot non-cooperative P.D. game.* In his own words,

> The difficulty in applying the ordinary theory of (Nash) equilibria to the one-shot case is that, except in the two person zero sum game, equilibria are not necessarily interchangeable. That is, if one player chooses in anticipation of getting one equilibrium and another chooses in anticipation of getting another, the resulting is not necessarily a third equilibrium.
>
> The same difficulty occurs in applying meta-game theory to the one-shot case, for this reason, the theory, like the theory of Nash equilibria, can only be applied to the one-shot case in the following rather unsatisfactory form: each player can work out the possible equilibria and decide to choose in anticipation of one of them; but whether that equilibrium results, and whether his choice turns out to be a rational one, depends on which equilibria the other players decide to anticipate. (Howard, 1966a, p. 185)

That meta-game theory may not determine a unique choice or that, in the absence of communication, there may be difficulty in coordination we do not regard as major drawbacks to Howard's approach. Our objection to applying Howard's approach to the one-shot non-cooperative case in that it does not solve the P.D. game but only bypasses it by making rather dubious assumptions about the psychological sets of the players.

Finitely Iterated Games

Let us now turn to finitely iterated Prisoner's Dilemma games played non-cooperatively some finite known number of times. Suppose that at each repetition the players, as usual, make their choices simultaneously and are informed of the outcome and receive payoffs resulting from each trial. Next, assume that each player's utility for the sequence of outcomes is the sum of his utilities in each of the component games.

Consider the iterated game as a "super-game," in which the moves are the component simple games. It has been shown (Luce and Raiffa 1957) that the only strategy of this super-game not dominated "in the wide sense" is the unconditional choice of strategy 2 throughout the iteration. Moreover, it has been shown that a strategy dominated in the wide sense cannot be one of an equilibrium pair; hence, the unconditional choice of strategy 2 throughout is the only equilibrium outcome of the finitely iterated P.D. game.[5]

The above result can be made intuitively clear by the following argument. Consider the rational choice on the n-th (last) iteration. Clearly, this iteration can be treated as a one-shot game. Hence, by our previous reasoning, the rational strategy on the n-th iteration is strategy 2 for both players. Now consider the (n-1)th iteration. Since the outcome on the n-th iteration was predetermined, the (n-1)th is now in strategic reality the last, and the same reasoning applies to it. This chain of backward reasoning leads to the inescapable conclusion that rational players will choose strategy 2 on each play of the iterated game. Yet, as Luce and Raiffa note, it is patently unreasonable to single out the repeated uniform choice of strategy 2 as the

5. Consider the P.D. game iterated n times. "Let us list the overall strategies of the super game as $A_1 = (r_1, r_2, \ldots r_{N_1})$, $B_1 = (s_1, s_2 \ldots s_{N_1})$, where N_1 is a finite but fantastically large number. One first notes that some of Row's strategies may be strictly dominated, i.e., there is at least one pair r_i and r_j such that r_i is never worse than r_j no matter what strategy Column employs and for some of Column's strategies it is better. Those dominated strategies may be thrown away with no loss to Row, leaving a new pure strategy set A_2 with $N_2 \leqslant N_1$ strategies. Similarly, Column has some strategies which are dominated relative to A_1 and these may be thrown away leaving a set B_2 with (because of the symmetry of the super game) N_2 strategies. Now, as long as Row knows that Column will confine himself to B_2 (Column would be stupid not to!) he may be able to throw away more strategies that are strictly dominated relative to any choice Column may make from B_2 (note: not from B_1). This results in a set A_3 with N_3 pure strategies, where $N_3 \leqslant N_2 \leqslant N_1$. Similarly, we define B_3. In this manner we go back and forth throwing away more and more strategies for Row and Column. A strategy which is thrown away at any stage of this process is said to be *dominated in the wide sense.*" (Luce and Raiffa 1957, pp. 99-100.)

solution of a P.D. game iterated some large but finite number of times. "It is not 'reasonable' in the sense that we predict that most intelligent people would not play accordingly. . . . [A]s long as our subjective a priori probability for our opponent's selection of strategy 2 is less than 1, we should not single out the unconditional iterated defection strategy." (Luce and Raiffa 1957, pp. 99-100, with some change of notation.)

Luce and Raiffa (op. cit., pp. 104-109) suggest two broad classes of strategies which they regard as preferable to the iterated unconditional choice of strategy 2. The first class involves the repeated use of the defection strategy only after its first use by the other player or only after some predetermined iteration, whichever comes first; otherwise, the use of the cooperative strategy. Note that if one player knew or had reason to believe that the other player would be choosing in accordance with some strategy from this class, the uniform iterated choice of strategy 2 would no longer be utility maximizing and might even be the worst possible strategy. In general, a utility-maximizing choice would "depend upon our subjective probability distribution over the set of strategy choices (assumed) available to our adversary" (op. cit., p. 104).

A class of strategies analogous to the two proposed by Luce and Raiffa has been suggested by Shubik. Although applied by him to infinitely iterated P.D. games with discount rates or stochastic stopping rules, they may be regarded also as "solutions" to finitely iterated P.D. games with or without discount rates (Shubik 1970). Shubik has been concerned with threat strategies. An example of a strategy that may be regarded as containing a threat in the iterated P.D. game is the following:

> "I will play my move 1 to begin with and will continue to do so, so long as my information shows that the other player has chosen his move 1. If my information tells me he has used move 2, then I will use move 2 for the immediate k successive periods, after which I will resume using move 1. If he uses his move 2 again after I have resumed using move 1, then I will switch to move 2 for the k + 1 immediately subsequent periods . . . and so on, increasing my retaliation by an extra period for each departure from the (cooperative) steady state." (Shubik 1970, pp. 187-88.)

Shubik requires that a threat strategy be both *plausible* and *sensible*. A threat strategy is said to be sensible if the steady state it advocates is Pareto optimal. The second concept, plausibility, is crucial to Shubik's treatment, but we shall not attempt to reproduce his rather lengthy analysis here (see Shubik 1959, pp. 222-32). Let us note, however, that for the threat strategy above, if $k \geqslant 1$, this strategy is sufficient to enforce (a, a) at least until the last iteration. Furthermore, according to Shubik, this threat strategy is extremely plausible.

Shubik remarks that "the least plausible threat is that which promises to punish the other player (and oneself) forever by playing move 2 every period in revenge for one violation of a desired steady state" (1970, p. 189). We regard it as useful to distinguish in the iterated P.D. game between (what we might loosely call) strategies of flexible response and strategies of *massive retaliation*. By the latter we mean one in which the defection counter-response, once initiated (or after having been initiated for the k-th time), becomes the strategy for the remainder of the iterations. By a strategy of *flexible response* we mean one in which the player never commits himself to lock in forever on the defection strategy. In general, we believe the latter type of strategy is more likely to prove plausible than the former and, more importantly, for the case of non-cooperative iterated games, less likely to commit the player to a pattern of behavior which he may come to regret.

"Solutions" to Infinitely Iterated Prisoner's Dilemma Games

All of the "solutions" discussed for the finitely iterated game apply also to the infinitely iterated game. The difficulty in treating the latter as a super-game is that in it at least some of the payoffs become infinite. One way of avoiding this difficulty is by introducing a discount factor c ($0 < c < 1$), so that the payoffs of the n-th iteration are multiplied by c^n, which assures the convergence of the infinitely summed payoffs of the component games (Shubik 1970). This approach we shall not discuss here.[6] Another way is to consider the average payoff per iteration associated with an infinite run. This is the approach taken by R. J. Aumann (cited in Shubik 1970, pp. 185-87). The average payoff per period will, of course, remain bounded and in the P.D. game can never exceed b.

> It can be argued that it is reasonable to assume that the players will consider their payoffs per period and will also take into account that, if they manage to achieve any temporary stationary state better than (d, d), this can be enforced as an equilibrium point. For if one player violates the stationary state then the other can choose his strategy 2 in every subsequent play, and, therefore, has a threat of punishment greater than the gain from violation. (Shubik 1970, p. 187)

Shubik's comments on this "solution" to the infinitely iterated game are ones with which we would concur. He notes that mathematically these results are impeccable, but goes on to demur that "the conversion of every outcome

6. We shall also not take up the question of infinitely iterated game with stochastic stop rules, i.e., such that after every single play "Nature" (a random factor) decides if the game is going to continue. If this random factor involves a fixed stop probability p, $0 < p < 1$; then it can be shown that the "solutions" are identical to those in the case of an infinitely iterated game with a constant discount rate $0 < c < 1$. In the one case, it is expected value that is shown to be finite; in the other, it is present value (see Shubik 1970, pp. 185-187, esp. Figure 5, p. 187).

We also shall not take up P.D. games played as games of survival, that is, games in which the play ends when an individual player is ruined (having assets less than some specified amount). Such survival problems can arise if we admit negative payoffs and assume each player begins with some fixed capital. In games of survival, the principal utility arises from staying alive, i.e., it is likely that utility is lexicographic. On the other hand, there might be individuals willing to "risk death" for some chance at a very large gain. Many of the interesting questions posed by P.D. games played as games of survival can be dealt with in the manner of Friedman-Savage in terms of the shape of the actor's utility function. Aspiration theory is also relevant (see Grofman 1972, chs. 3 and 4).

that satisfies individual rationality (i.e., gives each individual at least as much as he can guarantee for himself) into an equilibrium point leaves something to be desired and points to a weakness in the model itself" (Shubik, 1970, p. 187).

In an approach similar to that of Aumann, Overcast and Tullock (1971) have considered infinitely iterated P.D. games in which players adjust their percentage cooperative responses over a ten iteration span, in such a way as to avoid using dominated strategies, and have looked at some experimental data on choices in iterated P.D. games, but their predictions were not such as to be readily susceptible to statistical test.

The "solution" concept which we believe to be of the greatest promise is best represented in the work of Amnon Rapoport (1967) and his colleagues (Rapoport and A. Moshowtiz 1966; Rapoport and N. S. Cole 1968). Amnon Rapoport has treated a multiple-option form of expanded P.D. game as a decision problem under risk in which each player subjectively estimates the response contingencies of the other player on the basis of the outcomes of *all* previous iterations and chooses a strategy which is expected utility-maximizing; i.e., each player is assumed to act as a Bayesian decision-maker. However, his model is a rather complex one and we shall not discuss it further here. Instead, we shall turn to some Bayesian "solutions" to the infinitely iterated 2 X 2 P.D. game in which the decision problem is simplified by assuming that each player is using a Markovian strategy of a certain type and in which we look at the expected asymptotic level of payoffs associated with each strategy which might be chosen, in order to discover one which is "optimal" in the sense of maxmizing expected asymptotic payoff. In order to develop these notions more precisely, we must introduce some new terminology and notation.

Markov Decision Rules

Let us consider a symmetric two-person Prisoner's Dilemma game (cf. Figure 1) played over and over again with no change in the game matrix. In so doing, we generate a sequence of trials (game iterations, moves) with possible outcomes E_{11}, E_{12}, E_{21}, and E_{22}. A decision rule is essentially a maxim of conduct for a given player. Decision rules specify the probabilities of decisions, or combinations of decisions, as functions of the matrix entries, the outcomes of previous iterations of the game, and the number (n) of the present iteration.

We may formally define a *decision rule for player r* ($f_r^{(n)}$) as a function of n, of the matrix entries, and of the previous n - 1 outcomes of the game, which uniquely defines the probability of player r choosing alternative 1 at trial n. In symbols,

$$f_r^{(n)}(a,b,c,d,n,E^{(1)},E^{(2)},\ldots,E^{(n-1)}) \underset{df}{\equiv} p(r_1{}^{(n)}). \qquad (1)$$

A compound decision rule assigns probabilities to the possible outcomes at each iteration on the basis of the decision-rules used by each player. More formally, we define a *compound decision-rule*, $F_{E_{ij}}{}^{(n)}$, as a function of n, of the matrix entries, and of the previous n - 1 outcomes of the game, which uniquely defines the probability of outcome E_{ij} at trial n. In symbols,

$$F_{E_{ij}}{}^{(n)}(a,b,c,d,n,E^{(1)},E^{(2)},\ldots,E^{(n-1)}) \underset{df}{\equiv} p(E_{ij}{}^{(n)}). \qquad (2)$$

We define a *Markov decision rule* as a decision rule which is invariant with respect to the outcomes of the first n - k trials, for some k (i.e., which depends only on the last k - 1 outcomes), and which is applicable only to trial k and all later trials.

We define a *homogeneous decision rule* as a Markov decision rule which does not depend on n. For sequence of trials whose compound decision rule is homogeneous, knowledge of this rule and of the matrix entries will enable us to determine outcome transition probabilities invariant for all trials.

We define a *class m decision rule* as a homogeneous decision rule which is a function of the matrix entries and of the (n-m)th through (n-1)th trial outcomes and which is independent of n and of the outcomes on the first n - m - 1 trials. Thus a class 1 decision rule, $f_r^{(n)}$, for player r is a function of the matrix entries and of the (n-1)th trial outcome, which uniquely determines the probability of player r choosing alternative 1 at trial n and which is independent of n and of the outcomes of the first n - 2 trials. In symbols,

$$f_r^{(n)}(a,b,c,d,E^{(n-1)}) = p(r_1{}^{(n)}). \qquad (3)$$

Likewise, a class 0 decision rule (also called an *absolute decision rule*) is a decision-rule which is a function solely of the matrix entries. A class 0 decision rule does not depend on the outcomes of previous moves.

Similarly, a *class m compound decision rule*, $F_{E_{ij}}{}^{(n)}$, is a function of the matrix entries and of the (n-m) through (n-1)th trial outcomes which uniquely defines the probability of outcome E_{ij} at trial n, and which is independent both of n and of the outcomes of the first n - m - 1 trials. Thus a class 1 compound decision rule, $F_{E_{ij}}{}^{(n)}$, is a function of the matrix entries and of the (n-1)th trial outcome which uniquely defines the probability of outcome E_{ij} at trial n, and which is independent both of n and of the outcomes of the first n - 2 trials. In symbols,

$$F_{E_{ij}}{}^{(n)}(a,b,c,d,E^{(n-1)}) = p(E_{ij}{}^{(n)}). \qquad (4)$$

Such a decision rule generates a sequence of trials called a first-order Markov chain. For a sequence of trials whose compound decision rule is class 1, knowledge of this rule and of the matrix entries will enable us to determine outcome transition probabilities of the form P_{ij}, which we may represent in a stochastic matrix P, as in Figure 3 (Kemeny, Snell, and Thompson 1962).

Finally, let us define an *egocentric decision rule* as a decision rule for an individual player in which the only things relevant to player r's probability of choosing alternative 1 are his own payoffs and payoff entries. *Contextual decision rules* are decision rules in which the payoffs to and payoff entries of the other player(s) enter

$$\begin{bmatrix} p_{11,11} & p_{11,12} & p_{11,21} & p_{11,22} \\ p_{12,11} & p_{12,12} & p_{12,21} & p_{12,22} \\ p_{21,11} & p_{21,12} & p_{21,21} & p_{21,22} \\ p_{22,11} & p_{22,12} & p_{22,21} & p_{22,22} \end{bmatrix}$$

Fig. 3. Class 1 Transition Matrix[a].

a-This figure employs standard notation; i.e., $p_{hi,jk}$ is the probability that on any trial n the outcome will be E_{jk}, given that on trial n-1 the outcome was E_{hi}.

$$\begin{array}{c} \\ E_{11} \\ E_{12} \\ E_{21} \\ E_{22} \end{array} \begin{array}{c} \begin{matrix} \quad 1 & \qquad 2 \end{matrix} \\ \begin{bmatrix} y_{11} & 1 - y_{11} \\ y_{12} & 1 - y_{12} \\ y_{21} & 1 - y_{21} \\ y_{22} & 1 - y_{22} \end{bmatrix} \end{array}$$

Fig. 4. Column's Class 1 Decision-Rule.

$$\begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{matrix} E_{11} & E_{12} & E_{21} & E_{22} \end{matrix} \\ \begin{bmatrix} x_{11} & x_{12} & x_{21} & x_{22} \\ 1 - x_{11} & 1 - x_{12} & 1 - x_{21} & 1 - x_{22} \end{bmatrix} \end{array}$$

Fig. 5. Row's Class 1 Decision-Rule.

$$\begin{array}{c} \\ E_{11} \\ E_{12} \\ E_{21} \\ E_{22} \end{array} \begin{array}{c} \begin{matrix} E_{11} & E_{12} & E_{21} & E_{22} \end{matrix} \\ \begin{bmatrix} x_{11}y_{11} & x_{11}(1-y_{11}) & (1-x_{11})y_{11} & (1-x_{11})(1-y_{11}) \\ x_{12}y_{12} & x_{12}(1-y_{12}) & (1-x_{12})y_{12} & (1-x_{12})(1-y_{12}) \\ x_{21}y_{21} & x_{21}(1-y_{21}) & (1-x_{21})y_{21} & (1-x_{21})(1-y_{21}) \\ x_{22}y_{22} & x_{22}(1-y_{22}) & (1-x_{22})y_{22} & (1-x_{22})(1-y_{22}) \end{bmatrix} \end{array}$$

Fig. 6. Transition Matrix for a Class 1 Compound Decision-Rule.

into the calculation of the probability of player r's choosing alternative 1 on some trial n. The initial decision rules we shall consider are egocentric.

With these formal definitions out of the way, it is now appropriate to consider in less abstract terms what they imply. A class 0 decision rule is absolute in the sense that it is independent of any information about previous trial outcomes which might have been used to assign probabilities to an opponent's future moves. Maximin (Luce and Raiffa 1957), represented by the choice of alternative 2 in the game shown in Figure 1, is an example of a class 0 decision rule. In a strictly competitive (zero-sum) one-shot game, the maximin criterion provides the optimal decision rule against a rational player. In a non-strictly competitive game, the maximin strategy need not be optimal (op. cit.). Even in a strictly competitive game which is iterated, the minimax strategy may not be optimal against a player deviating consistently from "rational" strategy choices (Fox 1972).

A class 1 decision rule restricts itself solely to a consideration of the outcome of the (n-1)th trial and of the entries in the game matrix. Thus some information which might be useful in designing a strategy to maximize one's own utility is suppressed.

A class 0 decision rule is a special case of a class 1 decision rule, just as both are special cases of more general (class m) decision rules. Thus, if a minimax strategy is followed by both players in an iterated game, a sequence of trials which is a first-order Markov chain is indeed generated, but the associated stochastic matrix of transition probabilities will be characterized by identical rows.

In games where all decision rules are homogenous, the outcomes preceding the current one by more than a given number of trials, and the total number of trials that have taken place, do not affect the decisions of the players. Homogeneous decision rules do not permit of what is usually called learning behavior; thus the homogeneity restriction is an extremely important one.

We shall develop a notation for representing homogeneous decision rules. We may represent Column's (Row's) class 1 decision rule for an iterated two-person 2 X 2 game as in Figures 4 and 5. Utilizing a special form of matrix multiplication, we may represent the compound decision rule for this game as in Figure 6. It is easy to see that the matrix of Figure 6 satisfies the properties of a transition matrix (Kemeny, Snell, and Thompson 1962). For the balance of the paper we shall confine ourselves to class 1 homogeneous decision rules.

Bayesian Models

Markovian models have been used for the study of behavior in iterated Prisoner's Dilemma games, most notably by Anatol Rapoport (1966) and colleagues (Rapoport and Chammah 1965, Rapoport and Dale 1966). However, most of the Markovian models discussed by these authors are intended to be descriptive rather than normative. Besides Rapoport's own work, a considerable number of other studies deal with the effect on "cooperation" of changes in the stragegy used by the other player—where the experimenter (unbeknownst to the subject) assumes the role of the other player (see the comprehensive review in Oskamp 1970). While little or no theoretical rationale is customarily given as to why we might expect one strategy to be more successful than another in eliciting "cooperative" behavior, a considerable body of data has been built up. Unfortunately, very few of these studies have dealt with the effect of response-contingent strategies. Instead, they have usually used pre-programmed strategies in which the experimenter's choices are independent of the subject's responses, i.e., in which the experimenter utilizes a class 0 decision rule. (See, however, Suppes and Atkinson 1960, Overstreet and Pilisuk 1968). It is easy to see that, when one player uses a class 0 decision rule Prisoners' Dilemma, it is *always* optional for the other player to use the defection strategy. Thus such experiments are of quite limited interest for our purposes. With the exception of a tit-for-tat strategy, few class 1 or higher decision rules have been utilized by experimenters in P.D. game experiments.

For the P.D. game defined in Figure 1, consider Row facing Column who is using the homogeneous class 1 decision rule given by Figure 7. Such a decision-rule is a

$$\begin{array}{c} \\ E_{11} \\ E_{12} \\ E_{21} \\ E_{22} \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{bmatrix} p & 1-p \\ p & 1-p \\ 1-p & p \\ 1-p & p \end{bmatrix} \end{array}$$

Fig. 7. Partial Tit-for-Tat Decision-Rule.

partial tit-for-tat in which Row's (n-1)th move will be copied by Column on the nth play of the game with a probability p. Where p = 1, we have the full-fledged tit-for-tat strategy which has been used in a number of gaming experiments (see citations, Oskamp 1970). It seems well established that TFT strategy is superior in inducing cooperation; i.e., the choice of alternative 1, to any of the noncontingent strategies which have been studied (Oskamp 1970). Almost no experiments have been done with other contingent strategies. Exceptions are Bixenstine, Chambers, and Wilson 1964, Tedeschi et al. 1968, 1969. Only in the latter is there a comparison of the effect on cooperation levels of a 100% TFT strategy with other forms of contingent strategies. Unfortunately, none of the other contingent matching strategies used was of the form in Figure 7, thus a comparison between a 100% TFT and a p% (for varying levels of p) TFT strategy is impossible with existing data.[7]

Tit-for-tat and partial tit-for-tat strategies seem to us to be of special interest because of their simplicity and intuitive plausibility as strategic choices in an iterated P.D. game. In the remainder of this paper we shall try to show the implications of assuming one's opponent is using a strategy of the form specified in Figure 7.

If Column is assumed to be using a strategy given by Figure 7, what is the optimal strategy for Row to use in a P.D. game played an infinite number of times? It is easy to generate the conditions under which Row will prefer a pure strategy of 1 to a pure strategy of 2, and thus will prefer not to use his (iterated) defection strategy. Under the above assumptions as to Column's behavior, a Row player using the pure cooperation strategy given by Figure 8 will receive an expected asymptotic payoff which can be calculated by finding the steady state vector (eigenvector) for the (compound decision rule) matrix given by the product of the two strategy matrices obtained using the special product rule specified in Figure 6 (Kemeny, Snell, and Thompson 1962). This compound decision matrix is shown in Figure 9. The eigenvector for this matrix is simply

$$(p, 1\text{-}p, 0, 0). \qquad (5)$$

$$\begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cccc} E_{11} & E_{12} & E_{21} & E_{22} \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Fig. 8. Pure Cooperation Strategy.

$$\begin{array}{c} \\ E_{11} \\ E_{12} \\ E_{21} \\ E_{22} \end{array} \begin{array}{c} \begin{array}{cccc} E_{11} & E_{12} & E_{21} & E_{22} \end{array} \\ \begin{bmatrix} p & 1-p & 0 & 0 \\ p & 1-p & 0 & 0 \\ 1-p & p & 0 & 0 \\ 1-p & p & 0 & 0 \end{bmatrix} \end{array}$$

Fig. 9. Compound Decision-Rule for Pure Cooperation against Partial TFT.

The payoff to Row associated with this eigenvector is thus

$$ap + b(1\text{-}p). \qquad (6)$$

Similarly, the steady state vector for the matrix which gives the transition probabilities for a column player using the partial tit-for-tat strategy given in Figure 7 vs. a row player using a strategy of pure defection is simply (0, 0, 1-p, p)—the steady state row payoff for which is c(1-p) + pd. If

$$ap + b(1\text{-}p) > c(1\text{-}p) + pd, \qquad (7)$$

then, in an iterated P.D. game, a strategy of pure cooperation will be preferable to a strategy of pure defection (iterated minimax) against a player known to be using the strategy given by Figure 7.

But, for a P.D. game as defined in Figure 1, we may readily show that there always exists a p such that inequality (7) is satisfied, and that any such p must be > ½.

Lemma 1: For a, b, c, d satisfying the constraints given in Figure 1, there always exists a p, p > ½, such that (7) is satisfied.

Proofs of this and all subsequent lemmas and theorems may be found in Grofman and Pool (1975).

Lemma 1 tells us that there always exists a partial tit-for-tat strategy which can induce a rational opponent to prefer pure cooperation over pure defection (i.e., the choice of alternative 2, as required by an iterated minimax decision rule) and that it must involve more than a 50% level of reinforcement. (Obviously, a 50% TFT strategy is the same as random choice, and a less than 50% TFT strategy is one involving more "uncopying" than copying of the opponent's last move.) Note that in our proof we have used only the first of the two defining characteristics of the P.D. game in Figure 1.

In order to strengthen the above conclusions it would be useful to show, not just that pure cooperation is preferred to pure defection when inequality (7) holds, but also that the pure cooperation strategy is over-all optimal for p sufficiently large; i.e., that there is some p such that the strategy given by Figure 8 has the highest payoff of any class 1 strategy when used in iterated P.D. games against a player using the strategy given in Figure 7.

7. The authors are presently engaged in a series of experiments, at the Stiftung Rehabilitation (Heidelberg) and at the State University of New York at Stony Brook, on the impact of partial TFT strategies on cooperation levels. Some preliminary results are reported in Pool and Grofman (1975).

Theorem 1: There always exists p, $\frac{1}{2} < \frac{c-b}{a-b+c-d} < p \leqslant 1$, such that the strategy of pure cooperation yields at least as high a (asymptotic average) payoff as any other class 1 decision rule in an iterated P.D. game against a player using a partial tit-for-tat strategy.

In proving Theorem 1, we have established that there always exists a class of partial tit-for-tat strategies which can induce cooperation in "rational" opponents and that such strategies must involve at least a 50% level of reinforcement. If a+d = b+c, a condition which many of the P.D. games used in the experimental literature have satisfied, then we may give a somewhat stronger result.

Theorem 2: If a+d = b+c, then for *any* p, $\frac{1}{2} < \frac{c-b}{a-b+c-d} < p < 1$, the strategy of pure cooperation yields at least as high an (asymptotic average) payoff as any other class 1 strategy in an iterated P.D. game against a player using a partial tit-for-tat strategy.

We next present some results analogous to those established above, as to the conditions under which pure defection (cf. Figure 1) will be optimal against an opponent in a P.D. game using a (partial) tit-for-tat strategy.

Lemma 2: For a, b, c, d satisfying the constraints given in Figure 1, there always exists a p, $p \geqslant \frac{1}{2}$, such that the reverse inequality of (7), i.e.,

$$ap + b(1-p) < c(1-p) + pd, \tag{8}$$

is satisfied.

Theorem 3: There always exists a p, $\frac{c-b}{a-b+c-d} > p \geqslant \frac{1}{2}$, such that the strategy of pure defection yields at least as high (asymptotic average) payoff as any other class 1 decision rule in an iterated P.D. game against a player using a partial tit-for-tat strategy.

Theorem 4: If a+d = b+c, then for any p, $\frac{c-b}{a-b+c-d} > p > 0$, the strategy of pure defection yields at least as high a (asymptotic average) payoff as any other class 1 strategy in an iterated P.D. game against a player using a partial tit-for-tat strategy.

Theorems 2 and 4 establish that when a+d = b+c, a strategy of pure cooperation is optimal for $p > \frac{c-b}{a-b+c-d}$ and a strategy of pure defection is optimal for $p < \frac{c-b}{a-b+c-d}$ against a player in an iterated P.D. game known to be using a partial tit-for-tat strategy. Hence, *a row player in an iterated P.D. game who knows that a column player is using a strategy of the form given in Figure 6 need only consider two strategies: pure cooperation and pure defection.* In such a situation, Row should begin with whichever of his two strategies he on a priori grounds believes to be optimal against Column's unknown p level, and should then continue that strategy unless and until such time as the new information provided by Column's choices causes him to modify (Bayes's Theorem) his original judgment so as to conclude that the most likely p value of Column is such as to be optimally faced with his other pure strategy. Thus, in a P.D. game where a+b = b+c, against an opponent known to be using a (partial) tit-for-tat strategy, a player's choice of an optimal (class 1) decision rule is remarkably simple—only two decision rules need to be considered. Whether or not players are capable of (intuitively) making such simplifying calculations in experimental situations where they are told that the other player is using a tit-for-tat stragegy (with unknown p level) remains a matter for empirical verification—an investigation which the present authors have begun (Pool and Grofman 1975). Studies done at the Max Planck Institute for Psychiatry in Munich, in which programmed strategies have been used for P.D. and Chicken games, have found that, when the computer is programmed with a class 1 (homogeneous) decision rule, a player's own response probabilities tend to be a function of only the computer's last move; whereas, when the computer utilizes a decision rule which depends upon earlier choices of the player, the player's responses also become functions of moves before the immediately previous one.[8]

A number of studies have found that the utility to a given player of a given outcome in a two-person game appears to be a function, not simply of the player's own payoff for that outcome, but also of the difference in payoff between his outcome and that of the other player (see Grofman 1975).

These studies suggest that the importance of relative gain maximization as a motive is increased when the absolute magnitude of the payoffs is small. In the P.D. game, an increase in the weight attached to relative (as opposed to absolute) gain maximization must increase a player's (unconditional) probability of defecting, since it is only through defection that a player can establish a positive differential between his own and the other player's winnings. Thus we might anticipate that a concern for relative gain maximization could completely rule out a strategy of pure cooperation. It is possible, however, to show that, even if a player is concerned with relative as well as absolute gain maximization, there will still exist a (partial) tit-for-tat strategy (of the form given in Figure 7) such that use of that strategy by Column will compel a "rational" Row player to prefer a strategy of pure cooperation to a strategy of pure defection.

Theorem 5: Let K and 1 - K be the relative weights attached by a player in an iterated P.D. game to absolute gain maximization and relative gain maximization, respectively. If K = 0, then there exists no p such that pure

$$\begin{array}{c} \\ E_{11} \\ E_{12} \\ E_{21} \\ E_{22} \end{array}
\begin{array}{cccc} E_{11} & E_{12} & E_{21} & E_{22} \end{array}
\begin{bmatrix}
px_{11} & (1-p)x_{11} & p(1-x_{11}) & (1-p)(1-x_{11}) \\
px_{12} & (1-p)x_{12} & p(1-x_{12}) & (1-p)(1-x_{12}) \\
(1-p)x_{21} & px_{21} & (1-p)(1-x_{21}) & p(1-x_{21}) \\
(1-p)x_{22} & px_{22} & (1-p)(1-x_{22}) & p(1-x_{22})
\end{bmatrix}$$

Fig. 10. Transition Matrix when Column Plays Partial Tit-for-Tat.

8. Dirk Revenstorf, Max Planck Institute for Psychiatry, Munich, personal communication, June 28, 1973.

$$\begin{array}{c|cccc} & E_{11} & E_{12} & E_{21} & E_{22} \\ \hline 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 1 & 1 & 1 \end{array}$$

Fig. 11. Pure Defection Strategy.

cooperation is preferred to pure defection against a player using a p% partial tit-for-tat strategy. For all other K, there exists such a p, and it is given by

$$p > \frac{(2-K)(c-b)}{K(a-d) + (2-K)(c-b)} > \tfrac{1}{2}. \tag{9}$$

It is instructive to see how K influences the necessary p value for an actual P.D. matrix. For the P.D. game matrix given in Figure 12, if (as usual) K = 1, then the minimum p specified by (9) is 11/16. If K = 1/2, a $p > 33/38$ is required. If K = 1/4, then a $p > 77/82$ is needed. Thus, to be expected, as concern for absolute gain maximization is replaced with concern for relative gain maximization, a reinforcement level (p) nearer to 1 is needed to induce cooperation rather than defection in a "rational" player.

$$\begin{array}{c|cc} & 1 & 2 \\ \hline 1 & (5,5) & (-3,8) \\ 2 & (8,-3) & (0,0) \end{array}$$

Fig. 12. Representative P.D. Game.

CONCLUSIONS

In the literature, the two basic approaches to "solutions" to (noncooperative) P.D. games are (1) to regard either pure cooperation or pure defection or (2) to select a strategy which is utility-maximizing given (a) restriction of the realm of strategic choice to some class of decision rules, and (b) subjective probability assessments to one's opponent choice of strategies from among this class. If we look to asymptotic payoff, then we have established the existence of partial (p%) tit-for-tat strategies capable of inducing pure cooperation in a "rational" player cognizant of the strategy his opponent is using and unable to change it. Indeed, even when a player weights in his utility function both absolute and relative payoff, we have shown that pure cooperation will be preferred to pure defection for sufficiently large p.

We hope to have shown that even in iterated P.D. games, where the defection strategy has powerful attractions, repeated defection need not be the optimal response against certain other strategies. However, whether and to what extent partial tit-for-tat strategies will indeed induce cooperation is a matter for experimental investigation—an investigation which the present authors have recently started.

ACKNOWLEDGMENTS

The senior author is deeply indebted to Anatol Rapoport for very extensive editorial comments and suggested revisions of an earlier version of this paper, which have been incorporated here. Professor Rapoport should not be held responsible for the views about solution concepts expressed in this paper, or for the errors of interpretation or mathematical infelicities remaining in it.

Both authors wish to acknowledge the support given their research by the University of Mannheim and the SUNY (State University of New York) Research Foundation, the Stiftung Rehabilitation (Heidelberg), the Deutscher Akademischer Austauschdienst, the Council for European Studies, and the Alexander von Humboldt-Stiftung.

REFERENCES

AXELROD, R. "Conflict of Interest: An Axiomatic Approach," *J. Conflict Resolution*, 2 (1967), pp. 87-99.

BIXENSTEIN, V. E., CHAMBERS, N., and WILSON, K. V. "Effects of Asymmetry in Payoff on Behavior in a Two-Person, Non-Zero-Sum Game," *J. Conflict Resolution*, 8 (1964), pp. 151-59.

CUNNINGHAM, R. L. "Ethics and Game Theory: The Prisoner's Dilemma." *Papers on Non-Market Decision-Making*, 2 (1966).

FOX, JOHN. "The Learning of Strategies in a Simple Two-Person Zero-Sum Game without Saddlepoint." *Behav. Sci.* 17 (May 1972), pp. 300-08.

GROFMAN, BERNARD. "Rationality and Choice in the Context of Two-Person, Non-Cooperative Games." Unpublished Ph.D Dissertation, University of Chicago, 1972.

————. "The Prisoner's Dilemma: Paradox Reconsidered." *Frontiers of Economics*, 1 (forthcoming).

GROFMAN, BERNARD and POOL, JONATHAN. "Cooperation as an Optimizing Strategy in the Prisoner's Dilemma Game." Unpublished, State University of New York, Stony Brook, March 1975.

HELD, VIRGINIA. "Rationality and Social Value in Game Theoretical Analysis." *Ethics*, April 1966.

HOPKINS, J. F. "Game Theory and Generalization in Ethics." *Rev. of Politics*, October 1965.

HOWARD, NIGEL. "The Theory of Meta-games." *General Systems*, Vol. II (1966a), pp. 167-86.

————. "The Mathematics of Meta-games," *General Systems*, Vol. II (1966b), pp. 187-200.

————. *Games, Metagames and Rationality*. Cambridge, Mass.: MIT Press, 1970.

KEMENY, J., SNELL, L., and THOMPSON, J. *Introduction to Finite Mathematics*. Englewood Cliffs, N.J.: Prentice-Hall, 1962.

LUCE, R. D. and RAIFFA, H. *Games and Decisions*. New York: Wiley, 1957.

OSKAMP, STUART. "Effects of Programmed Initial Strategies in a Prisoner's Dilemma Game." *Psychonomic Science*, 19 (1970), pp. 195-96.

OVERCAST, H. EDWIN and TULLOCK, GORDON. "A Differential Approach to the Repeated Prisoner's Dilemma." *Theory and Decision*, 1 (June 1971), pp. 350-58.

OVERSTREET, EDWARD and PILISUK, MARC. "Simulation Models of Sequential Choices in the Prisoner's Dilemma." *Paper #210*, Institute for Research in the Behavioral, Economic, and Management Sciences, Krannert School of Industrial Administration, Purdue (Indiana) University, August 1968.

POOL, JONATHAN and GROFMAN, BERNARD. "Computer Programs as a Means of Efficiency and Control in Cross-Cultural Experimental Games." *Experimental Study of Politics*, (forthcoming).

RAPOPORT, AMNON. "Optimal Policies for the Prisoner's Dilemma." *Psychological Review*, 74 (1967), pp. 136-48.

RAPOPORT, AMNON and COLE, N. S. "Experimental Studies of Interdependent Mixed-Motive Games." *Behav. Science* 13 (1968), 189-204.

RAPOPORT, AMNON and MOSHOWTIZ, A. "Experimental Studies of Stochastic Models for the Prisoner's Dilemma," *Behav. Science*, 2 (1966), pp. 444-58.

RAPOPORT, ANATOL. *Fights, Games and Debates*. Ann Arbor: Univ. of Michigan Press, 1960.

————. *Strategy and Conscience*. New York: Harper & Row, 1964.

————. *Two-Person Game Theory*. Ann Arbor: Univ. of Michigan Press, 1966.

————. "Escape from Paradox." *Scientific American*, 217 (July 1967).

————. "Editorial Comment: Gaming Section" *J. Conflict Resolution*, 14 (June 1970).

RAPOPORT, ANATOL and CHAMMAH, A. *Prisoner's Dilemma*. Ann Arbor: Univ. of Michigan Press, 1965.

RAPOPORT, ANATOL and DALE, P. S. "Models for Prisoner's Dilemma." *J. Mathematical Psychology*, 3 (1966), pp. 269-86.

RAPOPORT, ANATOL and ORWANT, CAROL. "Experimental Games: a Review," *Behav. Science*, 7 (1962), pp. 1-37.

RUNCIMAN, J. C. and SEN, A. "Games, Justice and the General Will." *Mind* (October 1965).

SCHELLING, THOMAS. "Some Thoughts on the Relevance of Game Theory to the Analysis of Ethical Systems." In *Game Theory in the Behavioral Sciences*, Ira Buchler and Hugo Nutini, eds. Univ. of Pittsburgh Press, 1969, pp. 95-117.

SHUBIK, MARTIN. *Strategy and Market Structure*. New York: Wiley, 1959.

————. "Game Theory, Behavior, and the Paradox of the Prisoner's Dilemma." *J. Conflict Resolution*, 14 (1970), pp. 181-94.

SUPPES, PATRICK and ATKINSON, RICHARD C. *Markov Learning Models for Multi-person Interactions*. Stanford: Stanford Univ. Press, 1960.

TEDESCHI, J. T., ARONOFF, P., GAHAGAN, J., and HEISTER, D. "The Partial Reinforcement Effect and the Prisoner's Dilemma." *J. Soc. Psych.*, 75 (1968), pp. 209-15.

TEDESCHI, J. T., HEISTER, D. S., and GAHAGAN, J. P. "Trust and the Prisoner's Dilemma Game." *J. Soc. Psych.*, 79 (1969), pp. 43-50.

TERHUNE, K. W. "Motives, Situation and Interpersonal Conflict Within the Prisoner's Dilemma. *J. Personality and Social Psychology*, Sup. 8 (1968), pp. 1-24.

THOMPSON, G. "Game Theory and Social Value States." *Ethics* (October 1964).

WOLFF, R. P. "Reflections on Game Theory and the Nature of Value." *Ethics* (April 1962).

WRIGHTSMAN, LAWRENCE S., JR., O'CONNOR, JOHN, and BAKER, NORMA J. eds. *Cooperation and Competition: Readings on Mixed-Motive Games*. Belmont, Calif.: Wadsworth, 1972.