

Models for Prisoner's Dilemma¹

ANATOL RAPOPORT AND PHILIP DALE

Mental Health Research Institute, University of Michigan, Ann Arbor, Michigan

Five models of behavior in the Prisoner's Dilemma are evaluated with respect to two sets of data. Three of the models are relatively adequate in accounting for the observed time courses of outcomes, but are further differentiated by the variances. The "best-fitting" parameters for each model are used to suggest psychological hypotheses about the gross differences observed between male and female populations. The basic methodological problem remains of choosing a model which is accurate enough and at the same time both mathematically tractable and suggestive.

Prisoner's Dilemma is a two-person nonzero-sum game (non-negotiable) with payoff matrix,

| | | |
|-------|------------|------------|
| | C_2 | D_2 |
| C_1 | R_1, R_2 | S_1, T_2 |
| D_1 | T_1, S_2 | P_1, P_2 |

where $T_i > R_i > P_i > S_i$; $S_i + T_i < 2R_i$ ($i = 1, 2$). Strategy D dominates strategy C for both players. Hence "rational" play prescribes strategy D in a single play. This leads to the outcome (D_1D_2) , which is worse for both players than (C_1C_2) .

If the game is played n times in succession, the outcome of each play being announced, the sequence of plays can be represented as a supergame, in which the individual plays of the game become moves. The supergame does not in general have a strategy which dominates every other. It can be shown, however, that the strategy $D^{(n)}$, which prescribes D on every one of the n moves unconditionally, is the only equilibrium strategy and so, according to some definitions of "rational play" is "rational."

While strategy D in a single play can be defended on the strong grounds of absolute domination, it is difficult to defend the counter-intuitive strategy $D^{(n)}$ on other than purely formal grounds. But no other strategy can be singled out as "the" rational one. Hence the finitely iterated Prisoner's Dilemma game (where the number of

¹ The research work on which this paper is based was supported in part by National Institutes of Health grant NIH-MH-04238-05.

plays is known to both players) presents an even greater dilemma to normative game theory than the game played once.

Experiments with iterated Prisoner's Dilemma games yield fairly consistent patterns. When the number of plays is large (some hundreds), male university students (University of Michigan) generally begin with nearly equal frequencies of *C* and *D* choices. The first observable trend in massed data is a decrease in the frequencies of *C* choices. This trend continues for some 30-40 plays and is followed by a reversal toward greater frequency of *C* choices. Of 190 male pairs playing several variants of the game under several conditions, 130 were observed to have "locked in" on either (C_1C_2) or (D_1D_2) in the last 25 plays.² A grand average time course of *C* frequency is shown as the upper curve in Fig. 1.

The average time course of *C* is, of course, only a very small portion of the information to be gotten from the data. The data of the experiments is a collection of protocols, each 300 plays long, each play being denoted by one of the four outcomes (C_1C_2) , (C_1D_2) , (D_1C_2) , or (D_1D_2) . Besides the mean time course of *C*, an indefinite number of other statistics can be read from the data, for example, higher moments, distributions of lengths of runs, probabilities of the four types of outcomes conditional upon previous outcomes, etc. The frequencies of (CC) responses and of the combined "unilateral" responses $[(CD) + (DC)]$ are also shown in Fig. 1.

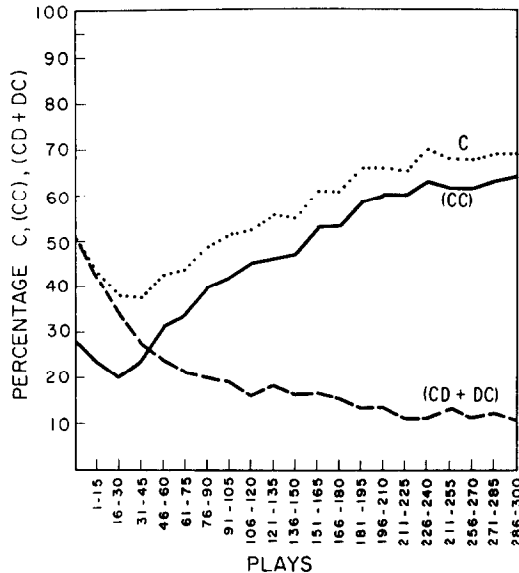


FIG. 1. Percentage of *C*, (CC) , and $(CD + DC)$ averaged over 190 male pairs playing eleven different games and 15 plays. The initial point is the average on the first play.

² A pair was counted "locked in" if 23 of the 25 last plays were either (CC) or (DD) .

An ideal mathematical model of the stochastic type would be one from which all the observed statistics could be derived as consequences. Such a model is too much to hope for. We begin by asking for considerably less, namely a stochastic model which would account for (1) the observed time courses (in massed data) of the frequencies of the four outcomes, and (2) the variance of the time-mean frequencies over the individual pairs.

We shall examine first the simplest stochastic model, namely a one-step Markov chain. As it turns out, this model will be shown to be clearly inadequate from the start, since it will not account even for the time course of the mean frequency of C .

Next, we shall examine four other models, three of which will be modifications of the one-step Markov model and one will be a stochastic learning model. These we shall examine with regard to the above-mentioned criteria and also with regard to comparing data from two populations, male and female.

MODEL 1. THE ONE-STEP (FOUR STATE) MARKOV CHAIN

Let (CC) be the probability of the outcome (state) (C_1C_2) ; (CD) of outcome (C_1D_2) , etc. We shall also refer to (CC) , (CD) , (DC) , and (DD) as states 1, 2, 3, and 4 respectively. The recursion equations of the probabilities are given by Eqs. 1

$$\begin{aligned}
 (CC)' &= \alpha_{11}(CC) + \alpha_{21}(CD) + \alpha_{31}(DC) + \alpha_{41}(DD) \\
 (CD)' &= \alpha_{12}(CC) + \alpha_{22}(CD) + \alpha_{32}(DC) + \alpha_{42}(DD) \\
 (DC)' &= \alpha_{13}(CC) + \alpha_{23}(CD) + \alpha_{33}(DC) + \alpha_{43}(DD) \\
 (DD)' &= \alpha_{14}(CC) + \alpha_{24}(CD) + \alpha_{34}(DC) + \alpha_{44}(DD)
 \end{aligned} \tag{1}$$

$$(CC) + (CD) + (DC) + (DD) = 1,$$

where if (CC) is the corresponding probability on play t , $(CC)'$ is the probability on play $t + 1$, etc. The α_{ij} are the transition probabilities between pairs of states.

The solution of the system of difference equations (1) gives each probability as a function of t . These functions will tend in the ergodic case toward asymptotic values obtained by setting $(CC)' = (CC)$, $(CD)' = (CD)$, etc., and solving the resulting linear system for (CC) , (CD) , etc., in terms of the α_{ij} .

Now the transition probabilities α_{ij} can be expressed in terms of the propensities of individuals to play C following each of the four outcomes (states). We shall call these the state-conditioned propensities. In particular, let

$$x_1 = p(C_1 | C_1C_2) \tag{2}$$

be the probability that following the outcome (C_1C_2) an arbitrarily selected individual labeled as player 1, will play C . Similarly,

$$y_1 = p(C_1 | C_1D_2) \quad (3)$$

$$z_1 = p(C_1 | D_1C_2) \quad (4)$$

$$w_1 = p(C_1 | D_1D_2). \quad (5)$$

The conditional probabilities x_2 , y_2 , z_2 , and w_2 are defined analogously, e.g., $y_2 = p(C_2 | C_2D_1)$, etc.

Consider some particular transition probability, say α_{23} . This is the transition probability from state (C_1D_2) to state (D_1C_2) . This transition occurred if player 1 played D following (C_1D_2) , while player 2 played C . If we assume that on a *particular* play, in the absence of communication, the players make their choices independently, the transition probability from (C_1D_2) or (D_1C_2) will be, in view of our definitions, $(1 - y_1)z_2$. Similarly, all the other α_{ij} can be defined in terms of x_1 , x_2 , y_1 , y_2 , etc. Consequently, the asymptotic solution of system (1) consists of four functions of the state-conditioned propensities x_i , y_i , z_i , w_i , and of the initial conditions.

The average values of the state-conditioned propensities and of the state frequencies can be taken as the observed frequencies of the corresponding choices, i.e., directly from the data, and so can the initial state probabilities.

A Markov chain with transition probabilities determined by the values of x_i , y_i , z_i , w_i and the initial conditions estimated from the data yields a time course of C which reaches its asymptote in about 30 plays. The general characteristic of this time course is rather insensitive to variations of the state-conditioned propensities and of the initial conditions. That is to say, although the asymptotic value of C clearly depends on the transition probabilities, the *number of plays required* to reach the asymptote within two decimal places does not exceed about thirty plays within a wide range of the transition probabilities. Thus, whatever values we assign to the state-conditioned propensities, the steady state will be reached much sooner than our data indicate. Moreover, the reversal of the trend from decreasing C to increasing (which is typically observed in the time course of C , as shown in Fig. 1) cannot be accounted for by the one-step Markov chain. We are therefore compelled to discard this model from the outset.

We shall next construct three generalized versions of the Markov chain and one stochastic learning model of the Bush-Mosteller-Estes type.

MODEL 2. MARKOV CHAIN WITH ABSORBING STATES

Following Cohen's conceptualization of iterated decisions (Cohen, 1963), we may assume that sometimes a player makes a permanent decision, e.g., "From now on I shall (or shall not) cooperate regardless of what happens." In our previous model, an

individual could be in one of two states; namely C , in which case he would choose C on the following play; or D , in which case he could choose D on the following play. In our present model, we assume that an individual can be in one of four states; namely,

- C : if in this state, he will cooperate on the next play but may defect later;
- D : if in this state, he will defect on the next play but may cooperate later;
- Γ : if in this state, he will always cooperate henceforth.
- Δ : if in this state, he will always defect henceforth.

The two-player system can now exist in one of 16 states. We can simplify the model somewhat if we make an additional not unreasonable assumption. Namely, passage into Γ by either player can be only from the (CC) state of the system, i.e., when both players have cooperated. Similarly, passage into Δ by either player can be made only from the (DD) state. This restriction excludes states $(\Gamma\Delta)$ and $(\Delta\Gamma)$, leaving 14 states in all.

Now we can no longer estimate all the state-conditioned propensities by simply counting frequencies of corresponding plays, because when a (CC) run continues to the end of the protocol, we do not know whether one or both of the players have passed into the Γ state (and if so, when), or whether they are continuing in the C state, and similarly for the DD runs. An estimation procedure based on a maximum likelihood method is tedious and is rather irrelevant to our purposes. Nor would such estimates be reliable in view of the number of free parameters involved and in view of the fact that even a protocol of 300 plays is not nearly long enough for this purpose. Instead, we shall simply seek the best fitting parameters by trial and error. We shall end our search when we have found a set of parameters which give visually estimated best fits of the time course of the four *behaviorally observable* outcomes, namely the double cooperative choices [which represent the union of (CC) , (ΓC) , and $(C\Gamma)$], the double defecting choices [which represent the union of (DD) , (ΔD) , and $(D\Delta)$], and the unilateral outcomes, which represent the remaining eight states. Having found the parameters which fit the time courses of the observable outcomes best, we put the model to a further test by examining the resulting variances.

Examination of massed data reveals that x increases with the number of preceding (CC) outcomes, also that w decreases as the number of preceding (DD) outcomes increases; i.e., that players are more prone to continue cooperating after they have cooperated for some time, and to defect after they have defected for some time.

It is important to note, however, that observations corroborating the above-mentioned conjecture do not warrant the conclusion that for any *individual* pair the probability of breaking a (CC) or a (DD) run decreases as the length of the run increases. For the observations may be also the result of a "natural selection" effect. Namely, the longest (CC) runs are likely to be observed in pairs whose members have the

largest x . In computing the conditional probability x as a function of the length of the (CC) run which precedes the play in question, one inevitably selects runs produced by players with larger x , as the runs become longer, since pairs with smaller x are not likely to produce the longer runs. Thus the apparent increase of x (or a decrease of w), as the length of the corresponding run increases, may be due to only a statistical, not a "psychological" reason. In the next model, this question will be raised. In the present model, however, the massed data are treated as if they were produced by pairs with identical parameters, or by a sort of composite of the pairs comprising the population. It is the fictional "psychology" of this composite pair which is postulated in this model; therefore it is legitimate to say that the x or the w of this composite pair is subject to a learning effect, even though the x 's of the individuals comprising the population may not be.

MODEL 3. THE FOUR-STATE MARKOV CHAIN WITH ONE STATE-CONDITIONED PROPENSITY SUBJECT TO LEARNING

This model is identical to Model 1 except that the parameter x is subject to "learning." Specifically,

$$x(t+1) = \alpha_x x(t) + 1 - \alpha_x, \quad (0 < \alpha_x < 1) \quad (6)$$

whenever (CC) occurs on the t th play, and $x(t+1) = x(t)$ otherwise.

Equation 6 implies that x will increase tending to 1. In principle, the other state-conditioned propensities could also be assumed to be subject to learning, tending toward 1, toward 0, or toward some intermediate asymptotic value λ , depending on which outcome obtained on the play in question. The most general model of this sort would contain 132 parameters, if all players were assumed to be identical.³ Here we are considering the simplest possible case with just one α referring to x and to the transition from (CC) to (CC).

MODEL 4. STATE-CONDITIONED PROPENSITIES DISTRIBUTED IN THE POPULATION

As we have said, the massed data may exhibit a progressive increase of x as the length of a (CC) run increases, while there is no such increase in the individual protocols. It can be shown mathematically (cf. Rapoport and Chammah, 1965) that if the x of each individual is independent of the length of the preceding (CC) run, and if the population is comprised of individuals in which the value of x is distributed in any manner whatsoever (assuming a variance greater than zero), then the massed data will necessarily show a positive "learning effect:" i.e., the observed x will increase with the length of the (CC) run.

³ Since each α refers to one of the 4 parameters and to one of 16 transitions, there are 64 α 's, similarly 64 λ 's, and 4 initial values of x , y , z , and w , assuming initial probabilities of the four states estimated from the data.

In the present model, we shall take into account the individual variations in $x, y, z,$ and w . Consider a population of pairs, each pair characterized by a set of values of $x, y, z,$ and w , so that each of these parameters has a given frequency distribution in the population. What will be the expected time course of the four states in the mass data produced by the population? In principle, the problem can be solved analytically as follows.

1. Solve system (1) to get the time courses of the four states as functions of $x_i, y_i, z_i,$ and w_i, t and of the initial conditions.
2. Assuming that $x_i, y_i, z_i,$ and w_i are independently distributed, with frequency distributions $\psi_{x_i}, \psi_{y_i},$ etc., multiply each time course by the product of the eight distributions and perform the eight-fold integration with respect to all of these variables. The resulting function of t and of the initial conditions will be the time course of the corresponding state averaged over the pairs, and so to be compared with the corresponding time course observed in the massed data.

The mathematical task, although straightforward, is extremely tedious and moreover requires that we make specific assumptions concerning the frequency distributions of $x, y, z,$ and w in the population. We shall instead resort to simulation as follows. For each individual we shall estimate $x, y, z,$ and w directly from the corresponding response frequencies in this protocol. We shall then attribute these estimated values to each pair of "stat-subjects" and produce a protocol resulting from a one-step Markov chain. The set of protocols so produced will constitute our simulated massed data. We shall then extract from these simulated data the statistics of interest to be compared with the statistics extracted from the real data.

MODEL 5. THE STOCHASTIC LEARNING MODEL

Following Bush and Mosteller (1955) and Estes (1950), we assume that at each play the probability of choosing C is a linear function of that probability on the preceding play, whereby the parameters of the transformation depend on the preceding outcome. Thus, designating, as before, by primes the probabilities on plays following a given play, we have

$$C' = \alpha_{CC}C + (1 - \alpha_{CC})\lambda_{CC}, \tag{7}$$

if the last outcome was CC ;

$$C' = \alpha_{CD}C + (1 - \alpha_{CD})\lambda_{CD}, \tag{8}$$

if the last outcome was $CD,$ etc.

We assume the players identical, i.e., each characterized by the same learning parameters $\alpha_{CC}, \lambda_{CC}, \alpha_{CD}, \lambda_{CD},$ etc. Nevertheless the different "realizations" of

the indicated stochastic process will, in general be different, because each outcome is only probabilistically determined, and a given sequence of outcomes (which may be different in different realizations) in turn determines the probability of the next outcome. Thus, unlike the simple Markov chain, the probability of an outcome of any given play depends on the entire preceding sequence of outcomes.

In each of the models (except Model 4) we shall have many degrees of freedom in fitting the theoretical curves to the observed, since we shall be treating our parameters as free parameters to be adjusted so as to give the best fit. Nevertheless we shall not take advantage of all the degrees of freedom. For example, in the stochastic model, we shall not consider all of the eight parameters free. Instead we shall assume $\lambda_{CC} = 1$. This is equivalent to assuming that if the (CC) outcome were to continue indefinitely, it would eventually become fixated. Similarly, we shall assume $\lambda_{CD} = \lambda_{DC} = 0$; that is to say, unilateral cooperation, if continued, would eventually be extinguished. However, we shall keep λ_{DD} as a free parameter on the ground that equally convincing arguments can be given for fixation and for extinction of the iterated (DD) outcome.

In Models 2 and 3, values for y and z were estimated directly from the data. In Model 4, there are no free parameters, since x , y , z , and w are all estimated from the data for each player.

In all the models, the initial play was selected at random with equal frequencies from (CC), (CD), (DC), and (DD), as this is a characteristic feature of all our data.

Having derived the theoretical time courses and the variances from our models, and having adjusted the free parameters, if any, so as to give approximately the best fits, we shall be in a position to see whether any of our models are clearly "better" (in the sense of giving a better fit) than others. However, this is not the question we are most interested in. As it will turn out, the three models which give reasonably good fits (Models 2, 3, and 5) do so about equally well. Even though Model 2 seems to have an edge over the others, a categorical decision in its favor can certainly not be made, since one cannot take any of the models seriously as a representation of "the" underlying reality which is, in all likelihood, much too complex to be represented by a mathematically tractable model. The models are only "as if" representations of reality.

The question which is central to this investigation is the following. *Assuming* a particular model to be a (drastically simplified) representation of the dynamics underlying the situation we have examined, what can we say about differences between the compared populations? In particular, we shall be comparing the performances of men and women. The gross results are clear. Women cooperate considerably less than men. We want to know why. But in attempting to answer this question, we shall be searching for the answer only in the parameters of an adequate model, not in any notions we may have about the psychological characteristics of college men and college women.

Our models allow us to look at the data with greater resolving power, as it were. For example, suppose we assume Model 2. Comparing the values of the parameters x , y , z , w , γ , and δ , we shall be able to say whether women cooperate less because they

tend to defect more from CC (which is the case if their x is smaller), or whether because they are less likely to make the irreversible decision to cooperate (which is the case if their γ is smaller), or whether because they are more likely to make the irreversible decision not to cooperate (which is the case if their δ is larger), or all put together. On the other hand, if we assume, for example, Model 5, we shall inquire whether women cooperate less because they are slower in learning not to defect from (CC) (which is the case if their α_{CC} is larger), or whether because they are quicker in unlearning to cooperate unilaterally (which is the case if their α_{CD} is smaller), etc.

In this way, the psychological conjectures which are suggested by our results will be at all times direct consequences of the models and never *ad hoc*.

SOURCES OF DATA

The data from which the time courses shown in Fig. 1 were constructed comprise the protocols of 11 different games (all of the Prisoner's Dilemma type but with different numerical payoffs). Therefore, included in the variances over the individual protocols are variances due to the differences in the payoffs. This portion of the total variance could be eliminated by confining the data to protocols involving a single game. However, this would drastically reduce the number of pairs in the population of subjects and so the statistical stability of the results. We have therefore resorted to a compromise. We have selected protocols drawn from games which are "nearly" alike. As an example, consider four games shown below.⁴

| | C_2 | D_2 | C_2 | D_2 | C_2 | D_2 | C_2 | D_2 |
|-------|---------|----------|-----------|----------|----------|----------|----------|----------|
| C_1 | 1, 1 | - 2, 2 | 5, 5 | - 10, 10 | 5, 5 | - 10, 10 | 1, 1 | - 10, 10 |
| D_1 | 2, - 2 | - 1, - 1 | 10, - 10 | - 5, - 5 | 10, - 10 | - 1, - 1 | 10, - 10 | - 5, - 5 |
| | Game IV | | Game XIII | | Game XI | | Game XII | |

Note that the payoffs of Game XIII are those of Game IV multiplied by a constant factor. If the payoffs represent utilities measured on an interval scale, the two games are identical. It does not follow, of course, that the games are *psychologically* identical, which is an empirical question. It turns out, however, that at least with respect to the observed frequencies of C responses, the two games are fairly close to each other. Thus the average frequency of C responses in a population of female pairs each playing 300 iterations of Game IV was observed to be 0.48, while in a population of 33 female pairs playing 300 iterations of Game XIII, the frequency of C responses was observed to be 0.39. Similarly, male students playing games IV, XI, and XII chose C with

⁴ The designation of the games follows their designation in our experiments.

frequencies 0.66, 0.63, 0.59, respectively. Some of the differences are significant, but compared with the range of C over the eleven games used in the experiments, the differences are not large. Thus the variance over the individual pairs is considerably reduced by our selection of similar games. We have recruited our subject populations from pairs playing games IV, XI, XII, and XIII, shown in Table 1.

TABLE 1
SUBJECT POPULATIONS

| Game | No. of male pairs | No. of female pairs |
|-------------------|-------------------|---------------------|
| IV | 20 | 10 |
| XI | 10 | 10 |
| XII | 19 | 10 |
| XIII ^a | 0 | 33 |
| Total | 49 | 63 |

^a Experiments with Game XIII were performed by Mr. Lynn Morehous in a separate study.

RESULTS

Figure 2 shows comparisons between observed time courses and those obtained from a simulated four-state Markov process (Model 1). The parameters were chosen to give a "best fit" in the sense of closest agreement between the over-all frequencies of the C responses. In the case of the male subjects, the simulated curve has an asymptote of about 0.65 which is reached by the 60th play. The observed asymptote is 0.75 and is reached only by about the 195th play. Similar discrepancies are observed with regard to the (CC) curves. The simulation asymptote of the unilateral response curve, on the other hand, is too high.

In the case of female subjects, agreement between simulated and observed curves is closer. On the other hand, the asymptotic values are practically identical to the initial values so that the model is practically a static one.

Comparing the values of x , y , z , and w , we find only the y 's substantially different, a result which contradicts estimates of these values made directly from data. When we compare these values estimated from large populations of college men and women, we find the largest differences between x 's and z 's, while the y 's are practically equal and the w 's nearly so in the two populations (Rapopart and Chammah, 1965). This result suggests that women are less responsive than men to their partners' cooperative choices but not more "vindictive" in response to their partners' defecting choices. This effect is not corroborated by the "best fitting" four-state Markov chain model, which is perhaps another argument for its inadequacy.

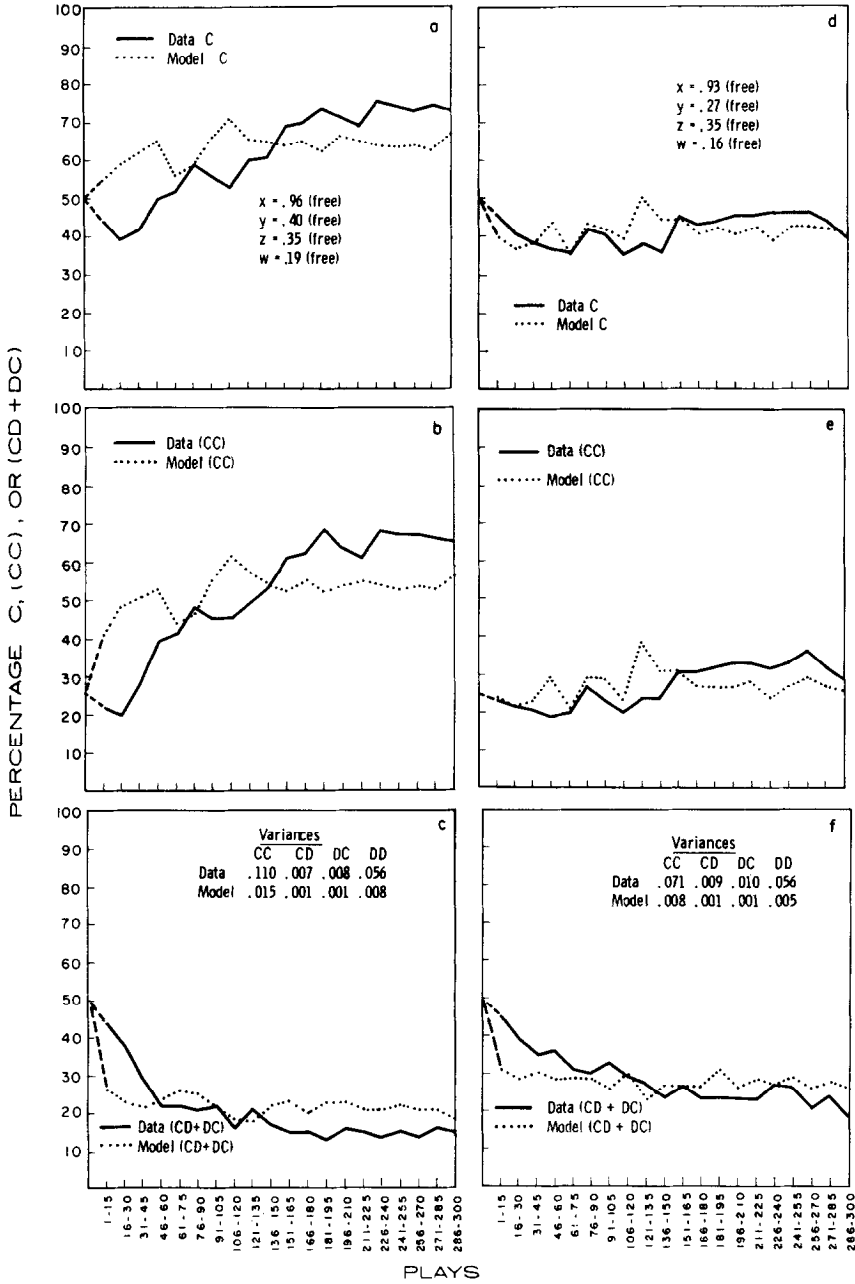


FIG. 2. Comparison of 49 male pairs and 49 pairs of simulated players, and 63 female pairs and 63 pairs of simulated players, for the Four-State Markov Chain Model. The initial point is the average on the first play. [a, b, c: Male; d, e, f: Female.]

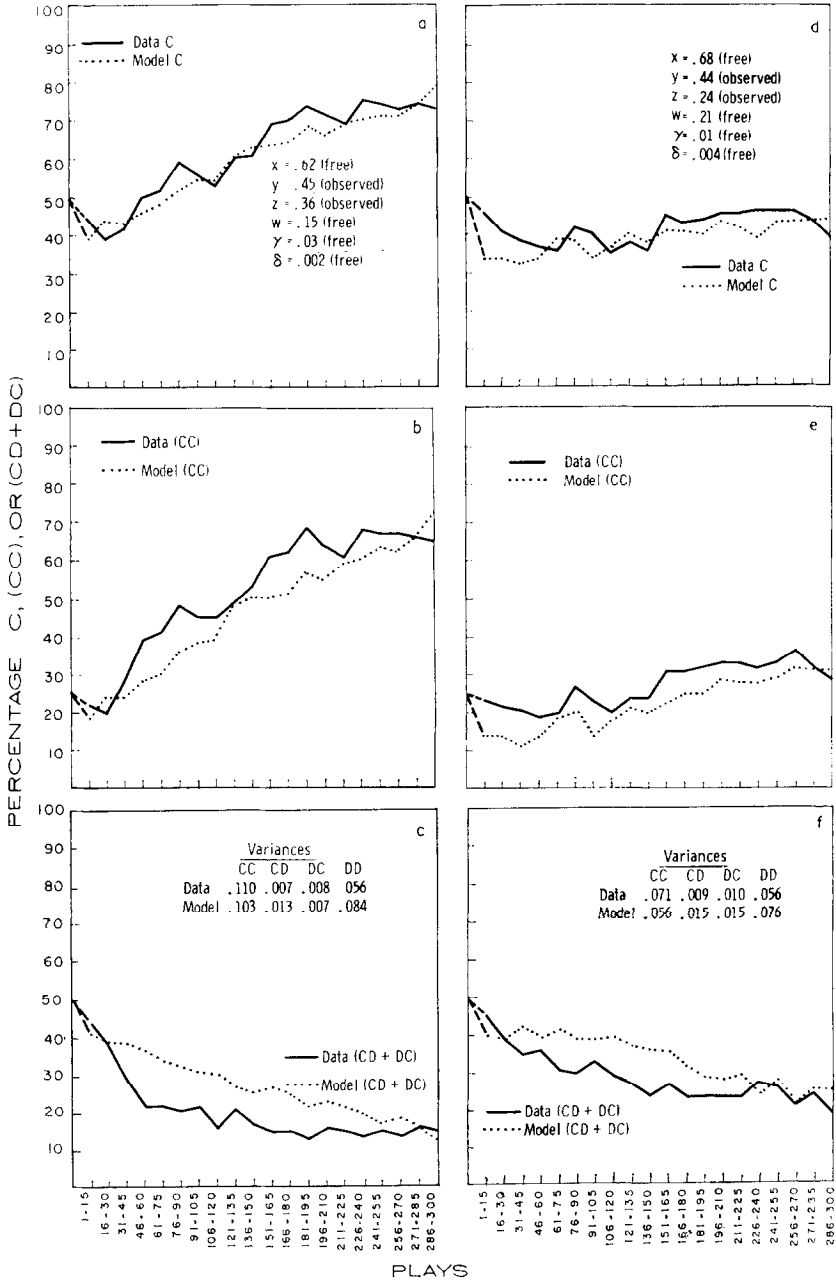


FIG. 3. Comparison of 49 male pairs and 49 pairs of simulated players, and 63 female pairs and 63 pairs of simulated players, for the Absorbing-State Markov Chain Model. The initial point is the average on the first play. [a, b, c: Male; d, e, f: Female.]

Comparing variances, we find that those derived from the model are much too low for both male and female subjects.

Figure 3 shows comparisons between observed time courses and those obtained from a simulated Markov chain with absorbing states (Model 2). As can be seen, the agreement of the C curves is quite good in the data from the males. The CC curve of the model, however, is too low, while the unilateral curve is too high. Similar results can be seen in the data from females. Comparing the parameters, we see the largest differences in favor of the males in z , γ , and δ . This suggests that (a) women are less responsive to the other's unilateral cooperation; (b) they are less prone to make the irreversible decision to cooperate; (c) they are more prone to make the irreversible decision not to cooperate.

The variances of the model and of the data are comparable.

Figure 4 shows comparison between observed time courses and those obtained from a simulated Markov chain model in which the parameter x is subject to learning, according to Eq. 6 (Model 3), while y , z , and w remain constant. The simulated C curve agrees with that obtained from the data rather well, except at the end where the simulated curve keeps climbing, as the model implies (since we took $\lambda = 1$). This discrepancy is even more pronounced in the (CC) curve. The simulated unilaterals, on the other hand, decrease too slowly. The comparison with the female data shows the same discrepancies.

Comparing the parameters, we see that the only prominent "cause" of the women's less cooperative performance is in a lower value of z . Their rate of learning with respect to x is only slightly slower than that of men, as is the initial value of x . Their y is only very slightly lower, while their w is actually slightly higher. Therefore, according to this model, the main difference between men and women seems to be a smaller propensity on the part of women to respond cooperatively to their partners' unilateral cooperation.

The variances from the simulated data are consistently too low both in men and in women.

Figure 5 shows comparisons between observed time courses and those obtained from a simulated stochastic learning model (Model 5). The C -curves agree very well, but again the simulated CC curves are too low while the unilateral curves are too high. In women, the simulated C -curve is too high. The difference between males and females is traced in this case to more rapid learning in males with regard to playing C in responses to CC , as shown by the smaller value of α_{CC} .

The variances are comparable, except in the case of DD , where the variances from simulated data are too low.

Figure 6 shows comparisons between observed time courses and those obtained from a four-state Markov chain model with x , y , z , and w estimated separately from the protocols of the individual subjects (Model 4). We see that the basic defects of the four-state Markov chain model are not eliminated. However, the agreement between the variances in the simulated and observed data are the closest in this case.

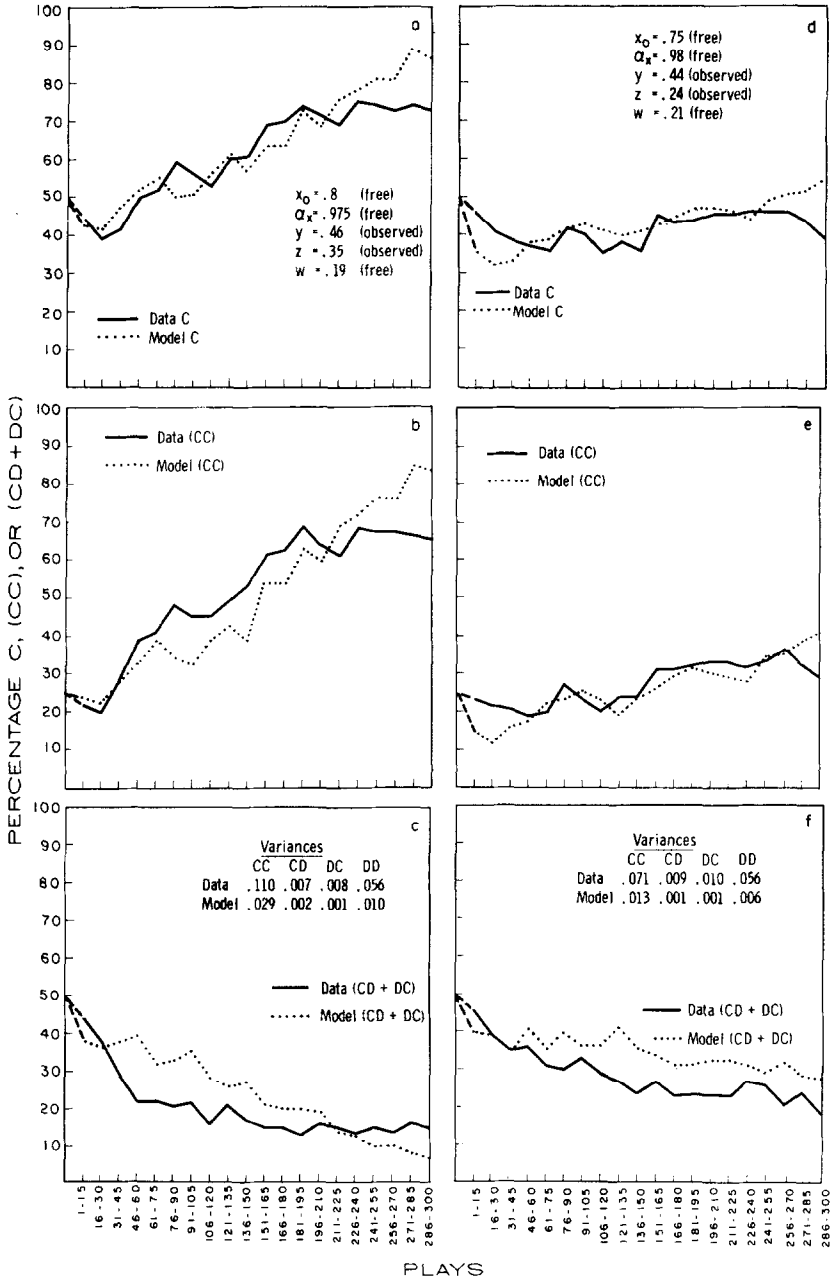


FIG. 4. Comparison of 49 male pairs and 49 pairs of simulated players, and 63 female pairs and 63 pairs of simulated players, for the Four-State Model with Learning on X . The initial point is the average on the first play. [a, b, c: Male; d, e, f: Female.]

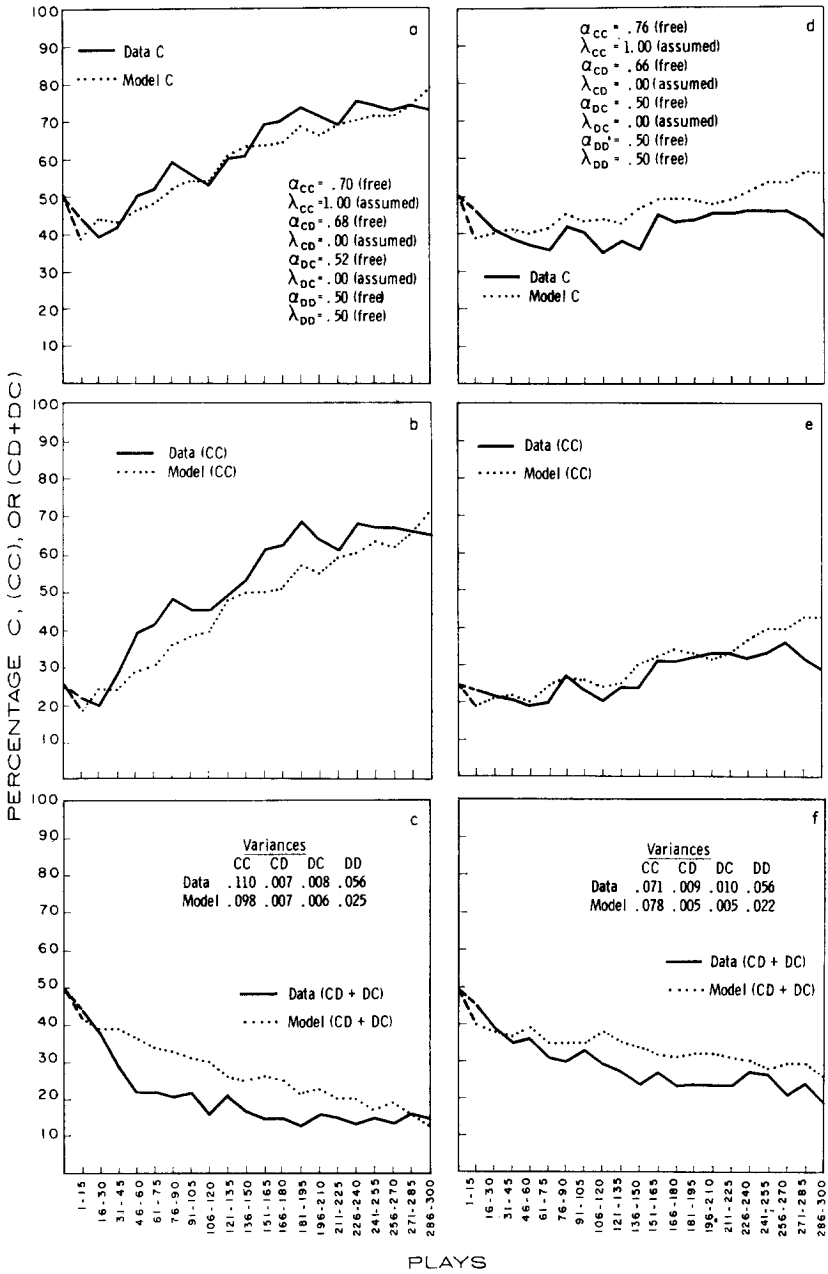


FIG. 5. Comparison of 49 male pairs and 49 pairs of simulated players, and 63 female pairs and 63 pairs of simulated players, for the Stochastic Learning Model. The initial point is the average on the first play. [a, b, c: Male; d, e, f: Female.]

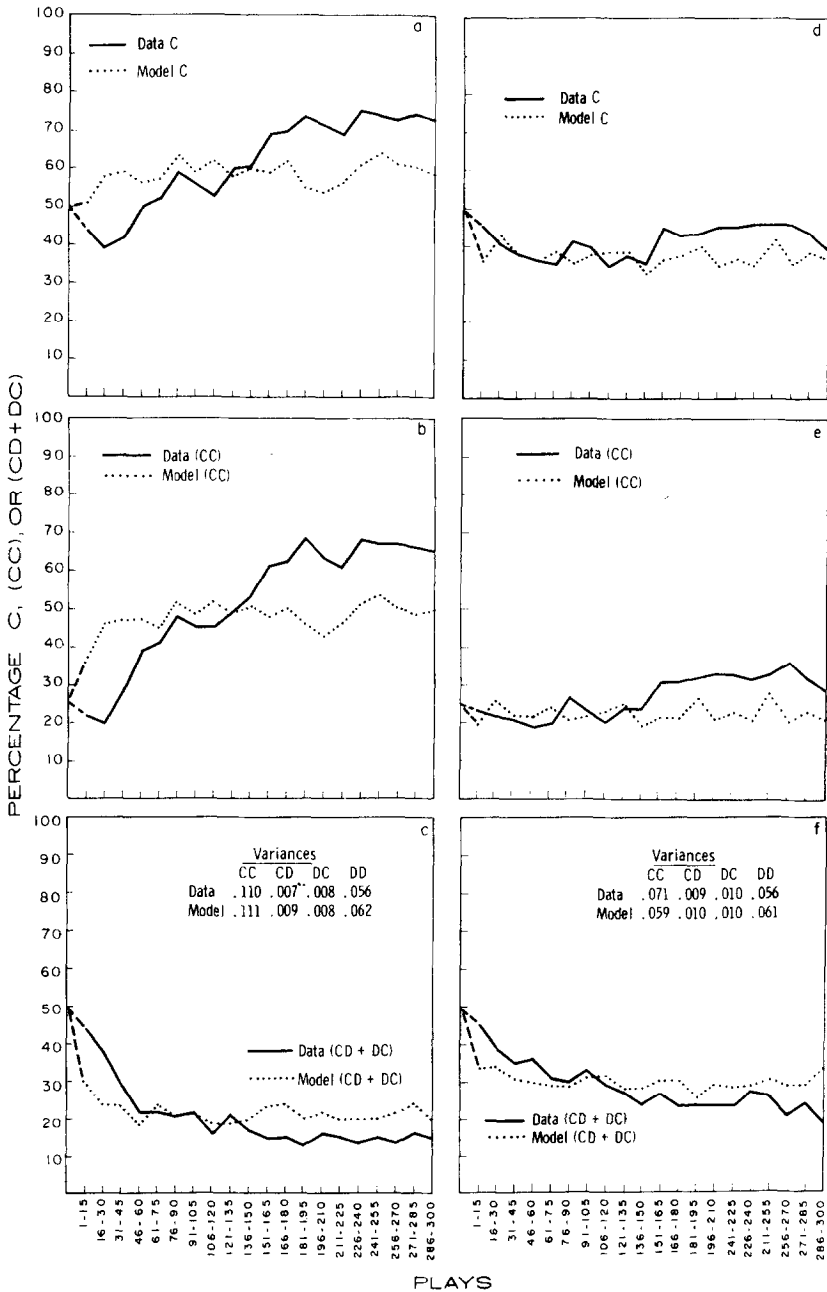


FIG. 6. Comparison of 49 male pairs and 49 pairs of simulated players, and 63 female pairs and 63 pairs of simulated players, for the Four-State Markov Chain Model with Distributed Parameters. The initial point is the average on the first play. [a, b, c: Male; d, e, f: Female.]

Here there were no free parameters. Accordingly we compare the estimated values of x , y , z , and w in the two populations shown in Table 2. Here we see that the difference in performance is traceable to two factors, namely differences in x and in z , while y and w are approximately the same in men and in women. This comparison corroborates the result obtained from the considerably large populations involving eleven different games. We interpret it to mean that while women are less prone than men to respond cooperatively to the other player's cooperation, they are not substantially more prone to respond with defection to the other's defection.

TABLE 2

ESTIMATED PARAMETERS

(Values of the state-conditioned propensities estimated from the protocols of 49 men and 63 women playing Games IV, XI, XII, and XIII (cf. p. 9))

| | Men | Women |
|-----|-----|-------|
| x | .85 | .72 |
| y | .45 | .44 |
| z | .36 | .24 |
| w | .21 | .19 |

CONCLUSION

We could probably obtain closer fits between simulated and observed time courses if we utilized all the degrees of freedom in choosing the parameters. On the other hand, the corroboration of our models would thereby become exceedingly weak, since almost any data can be fitted by appropriate models with a sufficient number of free parameters. Because of this "trade-off" and in the interest of simplifying the simulation procedure, we used only a few of our degrees of freedom (cf. p. 8).

Comparing the models, we feel that the four-state Markov chain models (Models 1 and 4) do not capture the dynamics of the process and must be discarded. Of the remaining models, Model 3 is the least satisfactory, largely because of the discrepancy in the variances. The choice is between Models 2 and 5. However, it must be kept in mind that Model 3 is the simplest of its kind. Its performance can probably be considerably improved by introducing additional learning parameters. One could also construct models in which the principles embodied in the models examined are combined, for example, both a learning process and absorbing states, or learning and a distribution of parameters in the population. The price of increased accuracy, however, in terms of the resulting complexity may be very heavy. The basic methodological problem remains of choosing a model which is accurate enough and at the same time

both mathematically tractable and suggestive. For some time to come, we expect, investigators will be guided in their choices more by intuitive or esthetic considerations than by compelling conclusions.

REFERENCES

- BUSH, R. R., AND MOSTELLER, F. *Stochastic Models for Learning*. New York: Wiley, 1955.
COHEN, B. P. *Conflict and conformity*. Cambridge, Mass.: M.I.T. Press, 1963.
ESTES, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94-107.
RAPOPORT, A., AND CHAMMAH, A. M. *Prisoner's dilemma*. Ann Arbor: The University of Michigan Press, 1965.

RECEIVED : July 20, 1965