Learning to Cooperate: Stochastic and Tacit Collusion in Social Exchange
Author(s): Michael W. Macy
Source: *American Journal of Sociology*, Vol. 97, No. 3 (Nov., 1991), pp. 808-843
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/2781785

Learning to Cooperate: Stochastic and Tacit Collusion in Social Exchange[1]

Michael W. Macy
Brandeis University

The Prisoner's Dilemma formalizes the social trap that arises when individually rational choices aggregate with mutually undesirable consequences. The game-theoretic solution centers on the opportunity for tacit collusion in repeated play. However, not all actors grasp the strategic implications of future interaction. Accordingly, this study reformulates the game as a stochastic learning model in which the behavior of interdependent actors is continually shaped by sanctions and cues generated by their interaction. Computer simulations of a two-person game show that adaptive actors are led into a social trap more readily than are fully rational actors, but they are also better at finding their way out. Prosocial norms appear to be a consequence rather than cause of cooperation but useful in promoting forgiveness of random deviance. The model is then elaborated as an N-way Prisoner's Dilemma. Simulations show how the effects of network size, density, mobility, and anonymity derive from a fundamental principle of collective action, that is, the need to reduce the number of choices that must be fortuitously coordinated in order to escape noncooperative equilibrium. The results also suggest how network structure might evolve in tandem with the cooperation it facilitates.

With the resurgence of sociological interest in the microfoundations of collective action, game theory has attracted the attention of sociologists working within a rational choice paradigm (Coleman 1986; Heckathorn 1988, 1989; Oliver 1984; Diekmann 1985). The point of departure in these studies is the paradox of social order among rational egoists whose pursuit of immediate gain leads to mutually suboptimal outcomes, or "social traps" (Platt 1973).

The problem of social traps may be formalized as a Prisoner's Dilemma in which the two players must choose between cooperation (C) or defection (D), a situation that generates a matrix with four possible outcomes

[1] This manuscript was written during a leave from Brandeis University under a Perlmutter Grant for 1989–90. I wish to thank Douglas Heckathorn, Michael Hechter, Peter Hedström, and anonymous reviewers for valuable suggestions and criticisms.

0002-9602/92/9703-0008$01.50

		Player 2: C	Player 2: D
Player 1:	C	$R = +1$ (Player 2); $R = +1$ (Player 1)	$T = +2$ (Player 2); $S = -2$ (Player 1)
	D	$S = -2$ (Player 2); $T = +2$ (Player 1)	$P = -1$ (Player 2); $P = -1$ (Player 1)

FIG. 1.—Payoff matrix for Prisoner's Dilemma

for the allocation of the payoffs—reward (R), temptation (T), sucker (S), and penalty (P)—as indicated in figure 1.

A typical illustration of this matrix is the exchange of resources for mutual gain (R) or to prevent a mutual loss (P). Each side stands to gain even more by taking more than they give (T), and risks being cheated by those who "free ride" on their willingness to sacrifice (S). The numbers ± 1 and ± 2 in figure 1 are arbitrary; so long as the payoffs conform to the inequality $T > R > P > S$, rational players will choose DD. The paradox is that they both would be better off picking C, yet whatever their opponent's choice, D offers the better payoff. If player 1 picks C, player 2 is better off picking D ($T = 2$) than C ($R = 1$), and if player 1 picks D, player 2 is still better off picking D ($P = -1$) than C ($S = 2$). Being rational, our two optimizers both pick D and receive P, the next-to-worst outcome for each. Rational choice theory thus provides a compelling account of how cooperation can suffer from what Hardin (1982, p. 6) calls "the back of the invisible hand."

This in turn poses a new puzzle: How do we explain the fact that social cooperation nevertheless seems to thrive? The rational choice solution requires that the game be ongoing, giving rational actors a *stake in the future*—an interest in avoiding acts that may incite retaliation as well as those that invite exploitation. An effective strategy for engineering a tacit collusion with a rational opponent must show that one is nice enough to be trusted but not so nice as to be taken advantage of (Axelrod 1984).

Those enlightened with sufficient foresight will thus appreciate both the futility of trying to exploit an equally sober interactant and the need to quickly retaliate against anyone tempted to test their resolve.

I. SOCIAL LEARNING THEORY AND RATIONAL CHOICE

The rational choice solution is elegant and appealing, particularly as it applies to collaboration between entrepreneurs, legislators, unionists, and other skilled negotiators. However, lay contestants may not always grasp why the single-play strategic solution to Prisoner's Dilemma does not obtain in an indefinite series (a game without a dominant strategy). They may be unwilling or unable to (1) estimate the probability of future engagements, (2) estimate the probability of being recognized by former (perhaps "faceless") interactants (3) appreciate the strategic implications of an ongoing interaction for inducing tacit collusion, and (4) place sufficient confidence in the capacity of their interactants to see the light. How might cooperation emerge among those whose self-interest is less prescient or who heavily discount an uncertain future?

There is an additional problem. When fitted with rationalist assumptions, game theory tends to reduce cooperation to the unintended by-product of the instrumental pursuit of private gain, an explanation that resonates very poorly with what we continually observe in social life: cooperation is typically accompanied by overt expressions of group solidarity. How might actors learn to cooperate not out of self-interest (however enlightened or "properly understood") but out of attachment to prosocial norms?

These limitations may explain the reluctance of sociology to share the enthusiasm for game theory shown by other disciplines. Still, this study is premised on the conviction that game theory is ideally suited to sociological concerns, although not in the way that is usually supposed. According to the conventional wisdom, "Underlying the entire structure of game theory is the key assumption that players in a game are *rational* (or *utility maximizers*)" (Zagare 1984, p. 7; emphasis in original). Although mixed-motive games can be useful for modeling strategic behavior, rational choice hardly exhausts the range of mechanisms by which payoffs might influence choices, as biological applications based on Darwinian theory clearly demonstrate (Maynard-Smith 1982). The game's "key assumption" is not rationality, it is instead what ought to be most compelling to sociology, *the interdependence of the actors*. The game paradigm obtains its theoretical leverage by modeling the social fabric as a matrix of interconnected contestants guided by outcomes of their interaction with others, where the actions of each depend on, as well as shape, the behavior of those with whom they are linked.

Strategic calculation is but one of the ways payoffs might shape behavior. Survival is another, with payoffs representing "fitness points" in natural selection (Axelrod 1984, pp. 88–107). Rapoport and Chammah (1965) have also proposed a behavioral model of Prisoner's Dilemma, which is based on social learning theory (see also Flood, Lendenmann, and Rapoport 1983; Gardner, Corbin, and Beltramo 1984). This study builds on their pioneering work. In a stochastic learning model of iterated Prisoner's Dilemma, the game payoffs operate as positive and negative sanctions or social cues generated by the interaction that guides the players toward anomie or solidarity through reinforcement and attenuation of cooperative propensities. The players are adaptive rather than purposive, "backward looking" rather than "forward looking," reactive rather than preemptive.

Although a stochastic learning model relaxes the strong rationality assumptions of conventional game theory, this need not imply that (1) human behavior is determined by stimuli, (2) that all behavior is adaptive, or (3) that all adaptive responses are unthinking. First, the model is not deterministic. Stochastic search introduces an element of uncertainty or "noise" that captures the idiosyncrasies of human behavior. This allowance for whim helps forgive the excessive frugality of the behavioral axioms. The model assumes only that the actors have a tendency to cooperate that is reinforced by the payoffs. They remain free to do the unexpected so long as their whims are randomly distributed. Random does not mean trivial; the central finding is that these idiosyncrasies are pivotal for escaping social traps.

Second, no claim is made that cooperation is never rational or that actors never exercise foresight. The point is not that actors always neglect the strategic implications of future interaction but only to ask what happens when they do.[2]

Third, social learning theory does not conflict with (nor does it share)

[2] Among economists, there is a growing "willingness to concede that the rationality assumptions of economic theory are not descriptive of the process by which decisions are reached and, further, that most decisions actually emerge from response repertoires developed over a period of time by what may broadly be termed 'adaptive' or learning processes" (Winter 1986, p. 244; see also Cross 1983). In the exchange of private goods, these processes tend to lead the actors, willy-nilly, into the equilibrium predicted for rational actors. Hence, market models need only assume that the actors behave *as if* they were rational. Elsewhere (Macy 1990*b*), I have questioned whether this "as if" principle is also available to rational choice models of the social dilemmas created by public goods, where opportunities to free ride, the lower impact of individual efforts, and the absence of competitive pressures on individual survival may trap adaptive actors in local maxima. The irony is that sociologists working with rational choice models may have no choice but to assume that actors in the social and political sphere *really are* forward-looking utility maximizers, which economists focusing on market behavior may not need to do.

the instrumentalist and egoistic assumptions of rational choice. Game-theoretic models of reinforcement and rationality share a common paradigm of payoff-conditioned behavior, and attention should be narrowly focused on the alternative specifications of this feedback, that is, on the relative weight of immediate and future outcomes on current choices. The instrumental basis of cooperative behavior is not the key point of contention.

To elaborate, adaptive actors chart their course on the fly, in response to changing signals, but this need not imply that cooperation emerges unintentionally, behind the backs of agents who remain unaware of what they are doing or why. Adaptive behavior can be manifested as consciously instrumental responses to the signals generated by choices, as in the children's guessing game of "getting hotter/getting colder." While the children's responses are deliberate and locally maximizing, their search strategy is unable to predict when they should take "one step backwards in order to take two steps forwards" (Elster 1979, p. 10). Myopic maximizers simply learn from their mistakes, repeating choices that seem to pay and avoiding those that prove costly. In cybernetic terms, they pursue their target relentlessly but cannot know its trajectory and therefore cannot intercept it by plotting a shortcut, a higher-order process that requires the capacity to anticipate its future moves. In short, "rationality is characterized by the capacity to relate to the future," which Elster (1979, p. viii) contrasts with "the myopic gradient-climbing" in adaptive or evolutionary processes.

Iterative Prisoner's Dilemma provides a convenient illustration of the distinction between backward-looking and forward-looking behavior. Fully rational action entails the singular capacity to recognize the logical implications of ongoing play. This capacity can be measured experimentally by observing mutually cooperative subjects when unexpectedly warned that the game is about to terminate. Rational actors, recognizing the implications of the endgame, should immediately stop cooperating despite having enjoyed considerable success up to that point (Luce and Raiffa 1957; for a more cautious assessment, see Hardin [1982]; for experimental evidence of "endgame" effects, see Tognoli [1975] and Murnighan and Roth [1983]). Backward-looking pragmatists will instead be eager to reap as much as possible from a demonstrably successful routine before the clock expires. They alter their behavior only as an aversive reaction to unwanted payoffs.

A. Instrumental and Normative Learning

Although learning theory can be used to model consciously instrumental (but myopic) behavior, it is typically applied to behavior that is unthinking or habitual. Adaptive responses in Prisoner's Dilemma need not be pragmatic but may instead be rule governed, with the payoffs sanctioning the legitimacy of the associated norms. The attachment to cooperative norms is thus strengthened when compliance entails the exchange of favors and when deviance is self-defeating. Conversely, the attachment declines when virtue is exploited with impunity.

Scott's (1971) theory of moral commitment shows how social learning theory can be applied to the "internalization of norms." Scott builds on Homans's behavioral model of social exchange in which "the behavior of one organism serves as a stimulus for the behavior of another . . . [which] in turn serves as a stimulus for the first" (Scott 1971, quote on p. 64; see also Homans 1961, p. 35). In complex social interactions, this mutual conditioning can generate patterns with varying degrees of regularity. Inconsistent sanctions lead to a cognitive response (based on expediency) that decays quickly. Instrumental compliance with social conventions depends on estimates of the magnitude and probability of sanctions, with no intrinsic attachment to the rule itself (Hechter 1987, pp. 3–5, 15–39). Hence, compliance ceases if the monitor leaves.

Over time, a consistent pattern of sanctioning may routinize habitual responses that then require only periodic maintenance. When norms are internalized, the action is no longer guided by the payoffs but instead becomes "an end in itself, regardless of its status as a means to any other end" (Parsons 1968, p. 75). While pragmatists may change tack after every wind shift, habits are slow to change. "When learning is established, extinction of the response by stopping the reinforcement is slow" (Boring 1950, p. 651), leading to compliance with a norm "at a spatial or temporal remove from its sanctions" (Scott 1971, pp. 88, 107).

To sum up, a stochastic learning model of Prisoner's Dilemma relaxes only the assumption that actors appreciate the strategic implications of future interaction. Like rational action, adaptive behavior can be fully conscious, deliberate, and expedient. However, while rational action is inherently instrumental, adaptive behavior can also derive from rote decision rules, unthinking habits and routines, and internalized norms.

The following example illustrates the difference between forward-looking and backward-looking behavior (whether instrumental or normative) in an everyday social interaction. Suppose two neighbors, Ego and Alter, stand to benefit by exchanging favors such as baby-sitting in a pinch, borrowing tools, getting a ride to work when the car is in the shop, watering the plants when the neighbors are on vacation, and avoiding aggravations such as playing loud music outdoors, stealing, or failing to mow the lawn regularly. These social exchanges provide each neighbor with greater security and convenience, yet each may at times feel annoyed by persistent requests, as well as tempted to take advantage of the other's goodwill. How might cooperation get started, and will it survive?

Let us begin by assuming Ego and Alter are fully rational and appreciate the opportunity for collusion. Although tempted to free ride, prudent foresight may lead them to restrain their selfish impulses and resentments and assist their neighbor, with each receiving R (see fig. 1). Each fears being made a sucker (S), and each is tempted to test the other's resolve and see how much they can get away with (T). However, they are chastened by their appreciation of what is at stake and their awareness that any such opportunism may invite retaliation (P) and undermine the confidence required for tacit collusion.

Now consider what happens if instead Ego and Alter remain instrumentally motivated and are invariably practical but neglect their stake in the future. Initially, our shortsighted neighbors may aggravate rather than aid one another (P), particularly if each thinks a neighborhood is nothing more than what one drives through to get home. However, Ego soon finds her behavior rather costly when emergencies arise and she needs a favor. Eventually, she may come to question the wisdom, if not also the propriety, of her self-serving conduct. If she gets suckered (S), she may conclude that neighborliness does not pay and revert to noncooperation, while the opportunity to take advantage of Ego with impunity encourages Alter to try again (T). However, should Alter happen to reciprocate Ego's gesture of goodwill (R), Ego senses that the new policy is working and continues. Over time, repeated reinforcement may lead to habitual compliance with the norm of "being a good neighbor" with little or no instrumental reflection. These civic obligations then motivate actions whose consequences strengthen attachment to the norm. With each exchange, the social ties across the back fence are strengthened.

These illustrations are not meant to suggest that adaptive behavior (whether instrumental or normative) will lead to more cooperation than that predicted for savvy strategists; indeed, there may well be less. The point is that the actors may not always have the temperament, cognitive skills, and information required to engineer a tacit collusion. Suppose instead they must grope their way toward a seemingly uncertain future, guided by the cues their actions generate. Learning theory does not resolve the social dilemma, it only reframes it: Where the penalty for cooperation is larger than the reward, and the reward for defection is larger than the penalty, how can penalty-aversive, reward-seeking actors elude the trap of mutual punishment?

B. Summary of the Findings

If players in iterative Prisoner's Dilemma fail to appreciate the strategic implications of future interaction, rational choice theory gives them little chance for escaping the trap. Computer simulations of a stochastic learning model confirm this. In Section II below, I elaborate my (Macy 1989) two-player model of social exchange by using simulations to draw out the implications of learning theory that are not apparent in the initial assumptions and to illustrate implications derived analytically.[3] The findings reveal a stable, noncooperative equilibrium into which even highly cooperative contestants are likely to gravitate. Where mutually contingent choices are conditioned by the sanctions they generate, the logic of interdependence tends to lead the players into this social trap.

The noteworthy finding, however, is that these adaptive actors can nevertheless escape, even though they fail to appreciate the possibility for collusion created by repeated play. Indeed, the model provides a convenient way to measure how much better they might do with a bit more foresight, and shows that it is a good deal less than what rational choice theory might lead us to expect. For nonstrategic players, cooperation depends not on the "shadow of the future," but on the salience of *immediate* outcomes. Those who respond quickly, without having to wait for long-term cumulative effects, can escape a suboptimal equilibrium through a fortuitous sequence of consecutive bilateral moves. Finally, the simulations suggest that attachment to cooperative norms is more likely to be a consequence than cause of successful collusion and that normative solidarity is useful primarily in promoting recovery from occasional deviance.

Section III extends the argument to a network of multiplex pairwise exchanges modeled as an N-way Prisoner's Dilemma. Computer simulations of adaptive behavior corroborate the main conclusions of rational choice theory on the difficulties of collective action in large, anonymous groups, but the model points to a very different diagnosis. The effects of network size, density, mobility, and anonymity are shown to derive from a fundamental principle of collective action: the need to reduce the number of choices that must be stochastically synchronized to escape anomic equilibrium. The results also suggest how network structure might evolve in tandem with the cooperation it facilitates.

[3] Simulated data can invite skepticism since they are not generated by experimental subjects. However, Turner (1987, p. 233) counters that simulation is "an important research activity, as it allows for the controlled analysis of more complex situations than is the case with laboratory experiments." Coleman (1986, p. 3) also argues that simulated games between "idealized persons" may be preferable to laboratory experiments where idiosyncratic factors introduced by "real persons" may obscure the theoretical implications, and Marwell, Oliver, and Prahl (1988, p. 503) show how simulations can be used to generate unexpected results.

II. A STOCHASTIC LEARNING MODEL OF SOCIAL EXCHANGE

The formal model assumes two contestants, each with some propensity to cooperate, p_{ij}, representing the probability that player j chooses cooperation at iteration i. Choices at each move are determined by the magnitude of p_{ij} relative to a random number n_{ij} from a uniform distribution, such that $C_{ij} = 1$ (player j cooperates at iteration i) if $p_{ij} \geq n_{ij}$, and $C_{ij} = 0$ if $p_{ij} < n_{ij}$. The choices of the two players then generate outcomes O_{ij}, as given by the payoff matrix for a Prisoner's Dilemma (R, S, P, or T in fig. 1), and the payoffs in turn modify their propensities to cooperate and defect.

The learning algorithm is an elaboration of a conventional Bush-Mosteller (1955) stochastic learning model for binary choice:

$$p_{i+1} = p_i + O(1 - p_i), \tag{1}$$

where p_i is the propensity to choose reinforced behavior at iteration i and O is a positive constant less than one. This model can be adapted to the Prisoner's Dilemma by allowing O to vary according to the relative magnitude of the payoffs. The propensity to cooperate is reinforced when cooperation is rewarded ($C_{ij} = 1$ and $O_{ij} > 0$) or defection is punished ($C_{ij} = 0$ and $O_{ij} < 0$):

$$p_{i+1,j} = p_{ij} + [O_{ij}(1 - p_{ij}^{(1/|O_{ij}|)})C_{ij}] - [O_{ij}(1 - p_{ij}^{(1/|O_{ij}|)})(1 - C_{ij})] \tag{2}$$

Since O can take on negative values, equation (1) was elaborated to allow for the reward for cooperation (C) as well as the punishment for defection ($1 - C$), as indicated in equation (2) by the two adjustments to p_{ij}, one positive and the other negative. Hence, the reward for cooperation is added to the propensity when $C_{ij} = 1$, while the negative payoff for defection is subtracted when $C_{ij} = O$, causing the propensity to cooperate to increase in either case. Conversely, if defection is rewarded or cooperation is punished, then the propensity to defect ($1 - p$) is reinforced, that is, $1 - p$ is substituted for p on both sides of equation (2) and $1 - C$ is substituted for C.

Equation (2) also assumes that reinforcement decays with the propensity, and that the larger the stimulus, the more rapid this decay, as given by the exponential expression $1/|O_{ij}|$. Hence, large rewards change behavior more than small ones, but 10 applications of a small reward have greater cumulative efficiency than five applications of a reward of twofold magnitude.[4]

[4] The Bush-Mosteller (1955) algorithm assumes that the exponent is fixed at unity and that O does not vary either. Here the model is applied to the Prisoner's Dilemma, and both the reinforcement and its rate of decay are allowed to vary with the payoffs. Freeing O while fixing the exponent alters equilibria but not the dynamics or the substantive conclusions; see Swistak (1990) and Macy (1990*a*).

It is useful to begin with the assumption that R and P reinforce the propensity to cooperate (R rewards cooperation and P punishes defection), while T and S decrease the propensity (T rewards unilateral defection and S punishes unilateral cooperation). Clearly, this constraint is not always appropriate. For example, aggravated neighbors may want nothing more from each other than an end to the loud noise, littering, and delinquent behavior of the children next door. Exchanges where Ego is penalized no matter what Alter does (with only a change in degree) will be introduced as a complication of the initial model. However, the analysis is focused at the outset on social dilemmas in which Ego is rewarded when Alter cooperates and penalized when Alter defects ($T > R > 0 > P > S$). For example, where the actors can choose between the exchange of gifts or blows, the breakdown of cooperation can lead to costly conflict, cease-fires can be a prelude to exchange for mutual gain, and benevolence may encourage aggression. Moreover, even a partial cease-fire may nevertheless be encouraging after a period of intense mutual conflict ($R > 0$). Similarly, the willingness to tolerate a neighbor's self-indulgent annoyance (with only token retaliation) may encourage its continuation ($T > 0$). Conversely, the failure to adequately share needed resources may prove costly ($P < 0$), while the failure to properly reciprocate a benevolent gesture may discourage its continuation ($S < 0$).

The payoff matrix can be reduced to two parameters, magnitude (σ) and severity (γ), where $\sigma = (R - P)/2$ and $\gamma = (T - S)/(R - P)$. It is also useful to impose the simplifying assumptions $R = -P$ and $T = -S$, giving $\sigma = R$ and $\gamma = T/R$. For example, a game with $\gamma = 2$ means that exploitation (CD or DC) is twice as rewarding as CC and twice as painful as DD. A magnitude of $\sigma = .01$ means that the actors learn very slowly, while $\sigma = 1$ means that the reinforcements are 100 times as large, such that the actors are likely to alter their course following each aversive outcome. Relatively low magnitude might correspond to a low interest in the outcomes, a noisy environment, or time lags that weaken the association between sanction and behavior. It can also model internalized normative responses that are extinguished more slowly than expedient responses based on conscious assessments of goal-attainment, as previously noted.

Since the four outcomes given by the payoff matrix follow automatically once we know the magnitude and severity of the sanctions, σ and γ inherit the constraints that define an iterated Prisoner's Dilemma ($T > R > P > S$ and $2R > T + S$, with the latter making the players better off cooperating than taking turns exploiting one another). Given the initial constraint $0 < |O| < 1$, and substituting $\sigma = R = -P$, $T = -S$, and $\gamma = T/R$, we obtain the a priori inequality $0 < \sigma\gamma < 1 < \gamma$.

To illustrate the learning algorithm, consider a game with $R = -P$

$= .1$ and $T = -S = .2$, or $\sigma = .1$ and $\gamma = 2$. Suppose the game begins with two rugged individualists, Ego and Alter. Their cooperative propensities at $i = 1$ are therefore $p_{1j} = 0$, leaving both certain to defect. This punishes both players ($O_{1j} = P = -.1$), causing their propensities to increase to $p_{2j} = .1$ (at $p_{1j} = 0$ there is no resistance to positive reinforcement). This process will reiterate until the pain is sufficiently unbearable (i.e., the repeated punishments cause p_{ij} to become sufficiently large) that Ego offers to cooperate. If Alter has yet to see the light, Alter is finally rewarded for defection ($O_{i2} = T = .2$) and becomes even less cooperative, while Ego is made a sucker with the same result ($O_{i1} = S = -.2$). Both propensities then drop up to 20 percentage points, with the downward reinforcements rapidly decaying as propensities approach their lower limit. If, on the other hand, Alter instead happens to also cooperate, both players find that they are rewarded for their benevolent gesture ($O_{ij} = R = .1$) and become even more likely to cooperate again. Should this occur, their cooperative propensities will increase by about .1 with each reinforcement until they become nearly 90% certain to cooperate, after which the rate of learning will drop off very rapidly as the players lock in mutually reinforcing cooperation.

A. Trapped in a Punitive Equilibrium

Computer simulations show how the learning process generates multiple equilibria in a Prisoner's Dilemma, one of which is both stable and punitive—a social trap from which the contestants may find it difficult to escape. The latent structural properties of the model are most clearly manifested by making the number of iterations required to alter behavior unrealistically large so as to minimize random disturbance of the equilibria. Let $\sigma = .01$ and $\gamma = 2$, giving a payoff matrix $R = -P = .01$ and $T = -S = .02$ for both players. In order to reduce unnecessary complexity, the paired players may be assumed to have identical initial propensities and to be subject to the same reinforcement schedule, which means their propensities can never diverge.[5]

Figure 2 reports the behavioral changes ensuing from three separate start values, chosen to illustrate equilibria. The three simulations reveal a punitive but stable equilibrium at .21, a threshold (or unstable equilibrium) at .79, and an absorbing state characterized by nearly 100% mutual cooperation.

[5] Macy (1989, p. 209) and Rapoport and Chammah (1965, p. 102) have shown that players with dissimilar initial propensities will promptly converge, with predominantly unilateral moves absorbed by *DD*. In the multilateral game, to be elaborated shortly, the players' propensities typically diverge.

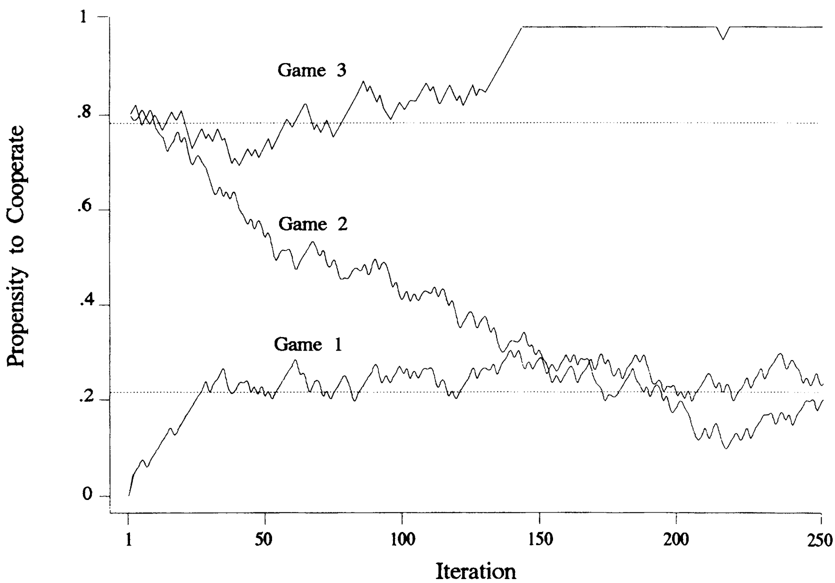

FIG. 2.—Equilibria for stochastic learning model of Prisoner's Dilemma ($\sigma = .01$, $\gamma = 2$; start values = 0, .8, .8).

At equilibrium, symmetric and asymmetric choices are distributed inversely to the ratio of their associated reinforcements. The symmetric moves CC and DD increase the propensity to cooperate, while asymmetric choices decrease it. Hence, equilibria occur where the probability of receiving an "upward" reinforcement times its magnitude (σ) equals the probability of receiving a "downward" reinforcement times its magnitude ($\sigma\gamma$).[6]

Since the magnitude of "upward" reinforcements is by definition less than that of the "downward," the stable equilibrium (p_e) must always be noncooperative ($p_e < .5$). A relationship at equilibrium will therefore always be predominantly punitive, entailing more incidents of mutual defection than cooperation (14 times as many in game 1 of fig. 2). This

[6] More precisely, the equilibrium propensity p_e may be obtained by

$$0 = 2p_e q_e \sigma\gamma(1 - q_e^{(1/\sigma\gamma)}) - (p_e^2 + q_e^2)\sigma(1 - p_e^{(1/\sigma)}), \tag{3}$$

where $q_e = 1 - p_e$, and σ and γ are the magnitude and severity of the payoffs as previously defined. Note that the probability of choosing C is p and the probability of choosing D is q. Hence the probability of either CC or DD is $p^2 + q^2$, and the associated payoff is σ. The reasoning for the asymmetric outcomes CD and DC is similar. What remains is the adjustment in the decay rate as propensities consolidate.

poses the central paradox of learning to cooperate in an iterated Prisoner's Dilemma. Although the players are attracted to R and recoil from P, even highly cooperative players respond to the reinforcements by gravitating toward a punitive equilibrium.

The noncooperative contestants in game 1 are caught in a self-defeating rut ($p_e = .21$). If both defect (or both cooperate), it makes the players less nasty, bringing them closer to the midpoint where the chances of unilateral defection are highest (unilateral moves are impossible when the mean propensity is 0 or 1 and most likely at .5). Unilateral moves then push both players back toward the noncooperative equilibrium. On the other hand, among cooperative players (those above the midpoint), the learning process is self-reinforcing. Above .5, mutual cooperation takes the players farther still from the treacherous midpoint. The Prisoner's Dilemma thus becomes a game of catch-22. Symbiotic behavior is self-reinforcing among players who are already cooperative and self-limiting among those who need to learn how.

Why then do the highly cooperative players at the start of game 2 not lock in mutual cooperation like those in game 3? While exploitation is self-limiting below .5, it too is self-reinforcing for cooperative players, setting into motion a series of recriminations. The contestants in game 3 are spared this fate only because their initial cooperative propensities are above the threshold value. For those above the threshold (or unstable equilibrium), cooperation is reciprocated sufficiently to offset the effects of expected betrayal, permitting a recovery of the lost ground through mutual cooperation before the disturbance can be repeated, as illustrated in game 3 at iteration 220. However, if either player is below the threshold propensity, cooperation invites sufficient incidents of exploitation that propensities will be ratcheted inexorably into anomic equilibrium, as happens in game 2.

Suppose exogenous sanctions are somehow imposed that make exploitation less rewarding but not altogether unattractive, that is, the game remains a Prisoner's Dilemma but one that is less severe. The model reveals a paradoxical effect. One might think that reducing the temptation to free ride would reduce the rate of exploitation at equilibrium. The unexpected result is that the incidence instead *increases*. Holding the other three payoffs constant, reducing T does indeed make the players less prone to defect and more cooperative than they would be otherwise. Lowering T thus raises the noncooperative equilibrium; at $T = R$, the game is no longer a Prisoner's Dilemma, and the equilibrium occurs at $p_e = .5$. But the closer the equilibrium comes to .5, the higher the equilibrium rate of exploitation, as previously noted. Hence, the less rewarding it is to exploit in a Prisoner's Dilemma, the more the players may be expected to do it.

B. Escaping the Social Trap: Those Who Hesitate Are Lost

By showing that the social trap is a stable, anomic equilibrium, the simulations also reveal a way out: *stochastic collusion.* Those unable or unwilling to engineer a tacit collusion may nevertheless enjoy an unintended facsimile if the reinforcements are sufficiently large. The problem in figure 2 is that the players alter their behavior too slowly, indeed, unrealistically so. More decisive actors can spring the trap with a short sequence of consecutive symmetric moves that gets them across the threshold and into "lock-in" before the law of averages catches up with them, as illustrated in figure 3. The parameters in this simulation are identical to those in figure 2 except that $\sigma = .1$. The larger payoffs mean that the players change their colors relatively quickly, without the need for long-term cumulative changes. (This also means that the reinforcements begin to decay earlier, raising the equilibria in fig. 3 to .26 and .81). It took about 20 iterations for the players to stumble into the coordinated sequence needed to cross the threshold and another 110 to attain lock-in after failing on the first attempt. This remains rather unrealistic for many if not most everyday relationships. However, as the magnitude increases further, the number of iterations needed to achieve lock-in declines exponentially, with lock-in requiring no more than two or three iterations if $\sigma \approx 1$, that is, if players always switch behaviors following aversive outcomes ($CC \rightarrow CC$, $DD \rightarrow CC$, and $CD \rightarrow DD \rightarrow CC$).

This solution to social dilemmas offers a clear alternative to rational choice explanations of the temporal logic of collective action. Among rational actors, the key to cooperation is the weight of the future relative to immediate outcomes. However, where contestants fail to enjoy the advantage of strategic foresight, simulations based on learning theory indicate that it is then the *weight of immediate outcomes that becomes critical.* Cooperation is thus unlikely where history casts a long shadow, that is, where routines can change only through long-term cumulative reinforcement. However, if both sides are prepared to shift course decisively, they can escape anomic equilibrium with stochastically synchronized moves. (In other words, the learning rates must be similar; if either side hesitates, both are lost.) Their behavior will look to all the world like that of clever strategists who have finally engineered a tacit collusion, but their collaboration is entirely fortuitous and unintended.

The need to reduce the number of choices that must be stochastically synchronized means that expediency may be more conducive to cooperation than is normative solidarity. Recall that the internalization of norms requires a consistent pattern of sanctions, after which the response will persist "at a spatial or temporal remove from its sanctions" (Scott 1971, p. 88). A series of successful exchanges will be needed to build attach-

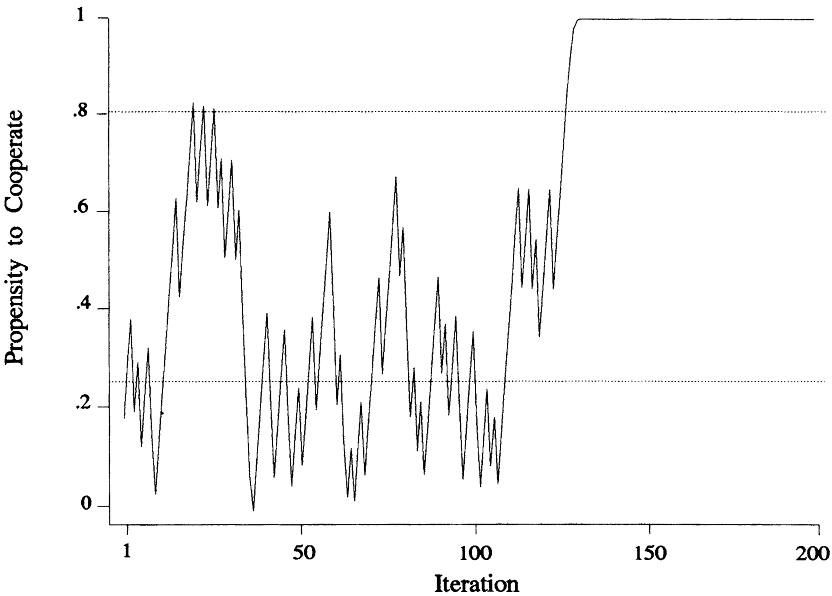

FIG. 3.—Locking in cooperation by stochastic collusion ($\sigma = .1$, $\gamma = 2$)

ments to cooperative decision rules among ambivalent actors. Noncooperative equilibrium thus tends to be anomic due to inconsistency in the application of sanctions: sometimes cooperation pays off and at other times it does not, depending on the choices of the interactor. Responses at equilibrium are therefore more likely to be pragmatic than normative. Since pragmatic responses are more decisive, they are more likely to facilitate stochastic collusion.

At the same time, expediency has a downside: the lack of forgiveness. Pragmatists may simply repeat responses that pay and may switch when they do not, thus promoting collusion but not forgiveness once lock-in is attained. While forgiveness is not needed to establish cooperation, it is necessary to prevent its collapse should one of the contestants deviate. Of course, cheating is highly unlikely at lock-in. Nevertheless, the game retains residual uncertainty—Selten's (1975) "trembling hand" that pushes the wrong button. "Overreaction" to random disturbance thus poses the danger that propensities may drop below the threshold for recovery. Hence, the model predicts that cooperation built on expediency will be less forgiving of occasional betrayal.

The need for forgiveness identifies the moment of internalized norms. Cooperation may be more easily established by pragmatists, but it is more easily preserved by those whose cooperative propensities are largely normative. Normative constraints may be expected to evolve only after the players lock in a consistent pattern of reinforcement. However, once normative solidarity is established, it is unlikely to collapse following a single betrayal since rule attachments change more slowly than do instrumental assessments of goal attainment. This suggests that normative solidarity develops only after cooperation has already been locked in, and that its province is not getting cooperation established so much as promoting recovery from occasional backsliding.

C. The Improvidence of Foresight

It remains to be shown how myopic pragmatists may sometimes learn to cooperate more readily than do rational strategists who exercise foresight. The behavior of the latter can be modeled as a stochastic process involving "response-conditioned" instead of "state-conditioned" propensities (Rapoport and Chammah 1965). Forward-looking players are influenced less by immediate payoffs (or states) and more by the responses of the interactant as these communicate the prospects for collusion or conflict.[7] So long as Ego cooperates, both models predict identical changes in Ego's behavior. If both Ego and Alter cooperate and Ego is backward looking, Ego is encouraged to continue because cooperation seems to pay (R). If Ego is forward looking, Ego is also encouraged to continue or risk the collapse of tacit collusion. By similar reasoning, both adaptive and fully rational actors will tend to retaliate when exploited (S). In short, mutual cooperation is self-sustaining whether the actors look backward or forward, and victims of exploitation may in both cases be expected to retaliate.

However, the models predict opposite reactions by Ego after Ego defects. Suppose a series of DD is interrupted by Alter's unilateral cooperation. Alter receives S and Ego gets T. Adaptive actors are encouraged by T to continue to defect. The victim, not the exploiter, is more likely to convert, making CD an unlikely transitional state between DD and CC. Rational strategists, on the other hand, resist the myopic temptation to free ride. Foresight cautions them to reciprocate rather than exploit

[7] Although strategic foresight has been emphasized here because of its prominence in game theory, response-conditioned strategies also apply to several other and quite different models that are behaviorally equivalent. The actors may learn to cooperate by imitating the behavior of interactants, as in role theory (Hechter 1987, pp. 67–68). Or propensities to cooperate may depend on feelings of warmth and hostility that reflect how one has been treated by the interactant. These diverse processes—strategic foresight, role modeling, and affective reciprocity—all motivate the actor to respond in kind to its interactants, converging in what is clearly a robust pattern of behavior, "tit for tat" (Axelrod 1984).

cooperative overtures by an equally sober interactant. Hence, to the extent that the players are forward looking, the impact of T will be attenuated or even reversed. Rational choice theory is thus much more optimistic that the free rider will convert, making CD a plausible transition to lock-in.

Rapoport and Chammah's classic experiments, based on iterated Prisoner's Dilemma, support this interpretation. Following a unilateral response, the defector appears to be much less likely to switch than does the sucker (1965, pp. 198–99). However, they found that "if the payoff matrix is displayed, the unilateral responses tend increasingly to become CC. . . . If the matrix is not displayed, on the other hand, this does not occur; the unilateral responses continue to be absorbed into DD responses. It appears that we have pinned down the role of the displayed matrix as a reminder to the subjects that a tacit collusion is possible" (1965, p. 95). Strategic foresight clearly seems to be helpful.

On the other hand, by reacting immediately to the sting of mutual loss, adaptive actors may have an advantage breaking out of a cycle of mutually self-defeating competition. Rational strategists know better than to indulge an opponent who continually defects. Prudence cautions each side to estimate the probability that a cooperative overture will be reciprocated before being the first to blink. Where each side expects the other to defect, each acts so as to reinforce the other's expectations. In classical game theory, DD is therefore a Nash equilibrium—a state in which strategic actors have no incentive to unilaterally alter their choices at the next move. Each side feins disinterest in cooperation, in the hope that the opponent will make the first overture, but when both play this game, the result can be a prolonged conflict. In stochastic representation, the downside of savvy behavior is that the impact of P is also attenuated or reversed. Reducing P lowers the noncooperative equilibrium; eliminating or reversing its impact creates a self-reinforcing absorbing state at $p_e = 0$, trapping the players in an endless series of mutual recriminations.

Figure 4 shows what happens in a game where the responses to both T and P are reversed. Cooperative propensities now increase with R and T but decrease with S and P, modeling a tendency to "tit for tat." Otherwise, the fixed parameters are identical to those for figure 3. The simulations show that the threshold of cooperation drops from .81 to .5, while the noncooperative equilibrium also drops from .26 to 0. The lowered threshold makes it more likely that initially cooperative players will lock in mutual cooperation. But, failing that, rational actors have much less chance of escape than do adaptive contestants more willing to alter their behavior in reaction to aversive cues. The shadow of the future clearly helps players consolidate a tacit collusion once cooperation has gained a foothold. However, in an anomic world, it is more liability than

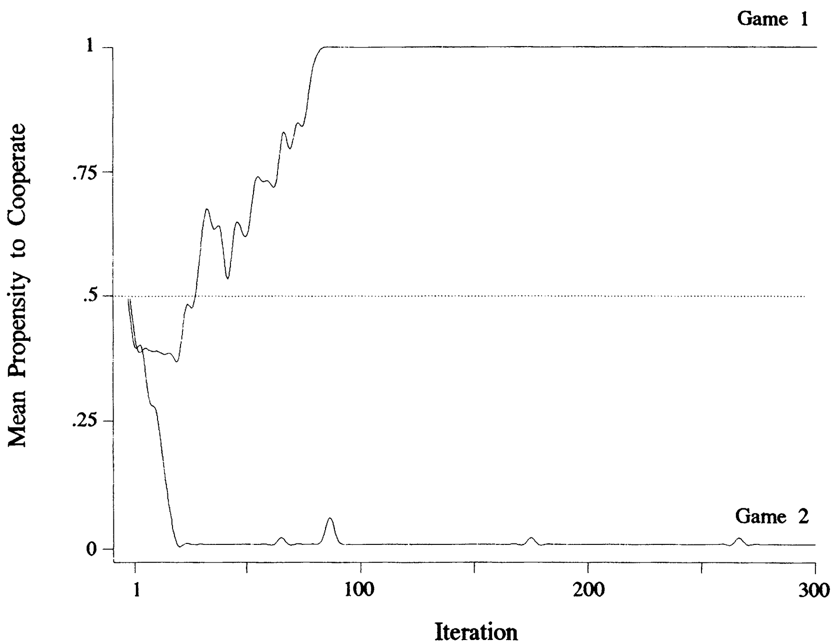

FIG. 4.—Equilibria with response-conditioned propensities ($\sigma - .1$, $\gamma = 2$; start values .5, .5).

asset. Adaptive players eventually escape equilibrium and lock in mutual cooperation, as illustrated in figure 3. Meanwhile, with identical payoffs, more sophisticated players remain trapped at $p_e = 0$, with no hope of escape.

Several experimental studies of iterated Prisoner's Dilemma show that mutual recrimination eventually wears down both sides. Rapoport and Chammah report a much higher initial incidence of *DD* than *CC* as "the 'hard realities' of the Prisoner's Dilemma game become impressed on the players. However, there is no percentage in *DD,* and sooner or later the players recognize this" (1965, p. 201). Other studies have also reported evidence of this pattern.[8] Simply put, anomic equilibrium does not appear to be the black hole depicted at the bottom of figure 4. Cooperation seems to gain a foothold in an asocial world through an aversive

[8] For studies of the effect of cognitive factors on successful collusion, see Lindskold and Finch (1981), Basu (1977), Pincus and Bixenstine (1977, 1979), Tognoli (1975), Murnighan and Roth (1983), and Steinfatt (1973). For the impact of noncognitive factors, see Carnevale, Pruitt, and Carrington (1982), Fox and Guyer (1978), Lacy (1978), and Clark (1983). For a brief discussion of differences between simulated and experimental results, see Macy (1990a).

reaction to punitive conflict, a response that neither requires nor benefits from "the capacity to relate to the future," Elster's benchmark for rational action. On that basis, let us proceed to games with more than two contestants.

III. COLLECTIVE ACTION AS AN N-WAY PRISONER'S DILEMMA

Collective action differs from cooperation in that it requires more than two actors. An N-way Prisoner's Dilemma models a specific type of collective action problem involving dyadic exchanges within social networks (Coleman 1987; Axelrod 1984). Most rational choice studies of collective action focus on the production of public goods in which outcomes are aggregated over the choices of the group as a whole (e.g., preserving the ozone layer, tithing to public television, or participating in social protest). The dilemma is that noncontributors receive the same benefit as volunteers, but if everyone free rides, no one benefits (Olson 1965; Hechter 1987; Oliver and Marwell 1988). The N-way Prisoner's Dilemma differs from the public goods problem in the following three ways:

1. Each player interacts with the group at large only by pairing directly with all members or indirectly with those who in turn pair with the others.
2. Each player's payoff depends only on the choices of a given partner, regardless of what others do. This contrasts with contributions to public goods where often "only a fraction of the benefits of one person's action accrues to that person" (Coleman 1987, p. 59).
3. Each player may choose to cooperate with some interactants and not with others. This differentiation of interactants is essential for the exercise of reciprocity necessary to secure a tacit collusion, a rational choice solution that is generally not available in the production of public goods.[9]

The earlier example of neighborly cooperation can be elaborated as an N-way Prisoner's Dilemma in which exchanges link each neighbor to others through a network of dyadic ties. Moreover, we need not assume that everyone relates equally to everyone else, and, as the size of the neighborhood increases, very different structures of interaction become

[9] Other rational choice solutions have been developed for the public goods problem, notably Olson's (1965) "selective incentives" and Hechter's (1987) systems for monitoring and sanctioning compliance with corporate obligations. Heckathorn (1989) shows how such a sanctioning system might be agreed upon by rational actors unwilling to cooperate in its absence, but not in games with only two players. Oliver and Marwell (1988) propose a "critical mass" solution based on the jointness of supply of most public goods. Elsewhere (Macy 1990*b*), I apply a stochastic learning model to the production of public goods, showing that the theory of critical mass does not require assumptions of rationality.

possible with different consequences for the emergence and flow of solidarity throughout the network. These patterns might be shaped by the physical location of houses, shared interests, social and cultural attributes, or other factors exogenous to the payoff matrix. They might also be altered by the tendency of the actors to be more attracted to cooperative neighbors, such that the structure of the network is both cause and consequence of successful cooperation. Multiple exchanges thus reframe the original two-player problem: Given what we have observed about the dynamics of cooperation in each dyad, how will group size and structure affect the emergence of collective action in N-player exchanges?

Rational choice theory predicts that collective action is less likely to succeed in large, anonymous groups where predators gamble that their victims will never see (or at least recognize) them in the future. Computer simulations based on a stochastic learning model corroborate the predicted effects of group size and structure but show that the results do not depend on the capacity of the actors to appreciate the strategic implications of social dilemmas. Hence failures may not be an incentive problem created by the discounting of future payoffs but rather a problem of timing and coordination. Finally, the model suggests how structures that are more conducive to stochastic collusion might evolve in response to the payoffs and how collective action, once established, might be sustained in the face of invasion by free riders.

A. The Problem of Group Size

Recall that rational actors will always cheat in single-play Prisoner's Dilemma since they do not expect to see their interactant again. The key to tacit collusion is the formation of stable, long-term relationships that carry the sobering expectation of future interaction. The fear of inciting retaliation gives rational actors an incentive to restrain their aggressive impulses.

Unfortunately, that restraint tends to dissipate in N-way exchanges. James Coleman uses a computer simulation of randomly paired interactions in iterative Prisoner's Dilemma to show how tacit collusion is compromised if free riders can become "lost in a sea of anonymous others" where they "anticipate getting away without retaliation" (Coleman 1986, pp. 66–67; see also Axelrod 1984, p. 49).[10] For example, despite the

[10] Note that in the production of public goods, group size may actually facilitate the formation of a critical mass needed for successful collective action, as argued by Hardin (1982) and more recently by Oliver and Marwell (1988). For a rational choice treatment of group-size effects, see also Olson (1965, p. 28), Raub and Voss (1986, p. 94), Opp (1986, p. 161), and Elster (1985, p. 354).

risk of collision, an impatient motorist may be unwilling to yield at an intersection to an anonymous driver with whom he is unlikely to interact in the future. Civility may thus be more common in insular, provincial settings not only because the actors identify with each other, share common values, or adhere to more altruistic norms but because fellow citizens cannot easily fade into the mass as they might in a highly rationalized metropolis. The greater the number of possible interactants, the lower the probability that any two players will interact again, and if they should, the lower the probability that the victim will recognize his or her predator. This leads each to discount the danger of retaliation and the promise of collusion. Hence the future casts a smaller shadow and collective action based on tacit collusion becomes less likely to succeed.

Computer simulations based on the stochastic learning model corroborate the pessimistic predictions of rational choice theory: cooperation is more likely to thrive where networks are composed of small, stable clusters among which there is little movement or interaction. However, the simulations show that the hypothesized effects of group size and structure do not depend on the assumption that the actors are able to appreciate the strategic implications of a social dilemma. As the number of interactants increases, each actor may find it difficult to prevent outcomes in one exchange from contaminating their propensities to cooperate with other contestants. If so, the players will alter their behavior toward one partner based in part on the outcomes of exchanges with another, making escape from noncooperative equilibrium especially difficult. The effect is equivalent to reducing the magnitude of the reinforcements in the earlier two-person game (changes in the decay rates not withstanding). Any players who remain in noncooperative equilibrium will tend to drag their interactants back down; hence, the more contestants, the more moves that must be stochastically coordinated to escape the social trap. Either everyone escapes or no one does. Whereas the probability of lock-in is about .015 in the two-player game depicted in figure 3 (based on eight consecutive bilateral moves beginning at noncooperative equilibrium), the probability plummets to .000225 (or $.015^2$) in an identical game but with twice as many players. Simply put, the players are doomed by the law of averages that is more strictly enforced as N increases, keeping the players trapped in a noncooperative equilibrium.

The analysis of the two-player game reveals a lower limit (at about $\sigma = .1$) below which stochastic collusion becomes highly impractical. By the same reasoning, there is also an upper bound on the size of an undifferentiated group in which collective action may be reasonably expected to emerge. By way of illustration, it is instructive to relax the assumption that $T > R$ and $\gamma > 1$, to consider a game at the margin of the Prisoner's Dilemma, with minimum severity and maximum reinforcement (σ and

γ approaching unity.) In this game, the players are likely to alter their response following aversive cues (S and P) and to repeat their response following gratifying outcomes (R and T). Hence, players will either be near-certain cooperators ($p_{ij} \approx 1$) or near-certain defectors ($p_{ij} \approx 0$). For cooperation to emerge, all players must interact only with those who are like-minded (which then makes everyone certain to cooperate). Two contestants will thus lock in cooperation within at most three iterations ($CD \rightarrow DD \rightarrow CC$). If $N = 4$, the players will attain lock-in within four iterations, assuming random pairings at each move, as demonstrated in figure 5.

As the size of the group increases, collective action by all members becomes more difficult. At $N = 20$, stochastic collusion is indeterminate but will usually occur within about 300 iterations. This is illustrated in figure 6, which reports the mean propensity of all 20 contestants in a game with $\sigma = \gamma = 1$. In undifferentiated groups larger than 20, collective action is unlikely to emerge, even with minimum severity and maximum reinforcement.

B. Jump-Starting Cooperation

Although theories of learning and rational choice converge in predicting the failure of collective action in large, anonymous groups, they point to very different remedies. There is a clear consensus in the literature on rational choice that collective action is likely to require the creation of a secondary sanctioning system if there are more than a handful of interactants (Olson 1965; Hechter 1987). Coleman's (1987) rational zealot is a case in point. Zealots allow themselves to be exploited by the cult leader

i=1		i=2		i=3		i=4
CC CC	→	**CC CC**				
DD DD	→	**CC CC**				
CC DD	→	**CC CC**				
CC CD	→	**CC DD**	→	**CC CC**		
	→	**CD CD**	→	**DD DD**	→	**CC CC**
CD DD	→	**DD CC**	→	**CC CC**		
	→	**CD CD**	→	**DD DD**	→	**CC CC**
CD CD	→	**DD DD**	→	**CC CC**		

FIG. 5.—Lock-in attained within four iterations with $N = 4$, $\sigma = 1$, and $\gamma = 1$.

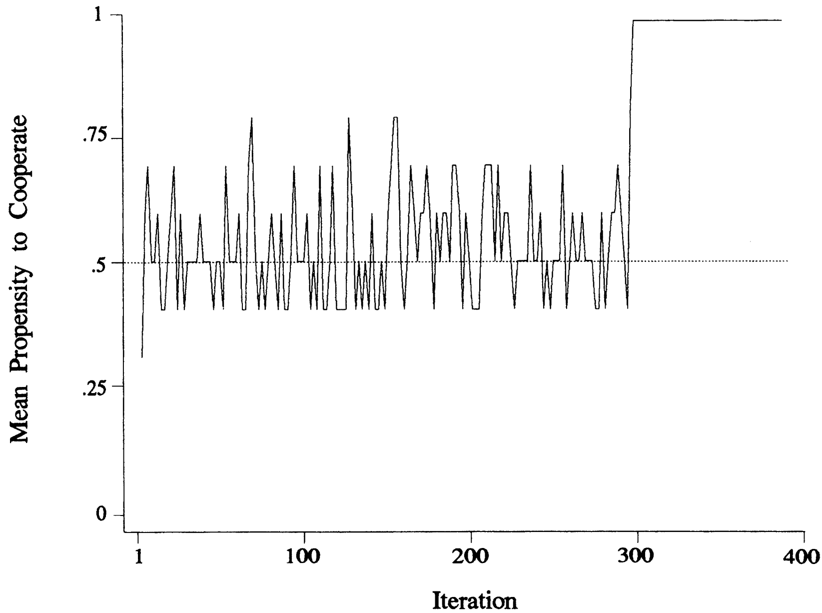

FIG. 6.—Stochastic collusion in 20-player game with maximum reinforcement and minimum severity ($\sigma = 1$, $\gamma = 1$).

in order to obtain the social approbation and status that is conferred only upon true believers. Since these psychic incentives are private (excludable) goods, they preclude free riding and thus transform the social dilemma into a conventional market transaction.

Learning theory suggests instead that the collective action problem may be mainly one of timing and coordination. Selective sanctions focused on exceptional participants may be doomed by the structural dynamics of the equilibria unless coupled with the pump-priming effects of public spectacle. Synchrony is the key to breaking out of social traps: short, intense mobilizations that command widespread attention so as to get everyone out of their noncooperative rut and over the threshold at the same time, at which point collective action becomes self-sustaining. Thus, periodic block parties and company picnics may jump start the exchange of resources and information (and a cease-fire of back stabbing and recrimination) among neighbors and colleagues. Similarly, annual religious gift-trading rituals are thought to facilitate more benevolent interpersonal relations (at least for a time).

Sanctions must not only be broadly applied but must also be appropriately timed, depending on the need for reward or punishment. In rational choice theory, "rewards and punishments do not differ and are interchangeable in their effect on individuals' rational decisions of whether to cooperate with collective action or not" (Oliver 1980, p. 1361, n. 7). In contrast, social learning theory shows that positive and negative sanctions perform entirely different catalytic functions. This can be illustrated by relaxing the assumption that $R > 0 > P$. Over time, ongoing conflict may acquire a sense of inevitability that numbs the players to the costs of mutual defection ($P = 0$). Or noncooperation might even be rewarding but suboptimal ($R > P > 0$). The stable equilibrium then approaches zero, precluding stochastic escape. In this anomic world, rewards to the occasional good citizen are counterproductive, indirectly encouraging those who prey upon benevolence. A shock to the system, deliberately induced by authorities or resulting from an unintended crisis, may be the only way to jolt the group out of its rut.

Once cooperation gains a foothold, the problem changes. Lock-in is induced by rewards for cooperation, not punishments for defection, and as the rate of cooperation increases, good citizens may be rewarded with less risk of feeding the sharks. This applies as well to games of conflict reduction, where $0 \geq R > P$. For example, punishing siblings for bothersome quarreling tends to be redundant.[11] Truces are not self-reinforcing without rewards for self-restraint and are therefore likely to collapse at the first unilateral defection. However, rewarding the contestants for a cease-fire may preserve it long enough for mutually beneficial, self-sustaining forms of interaction to emerge.

C. Network Structure and Stochastic Collusion

So far we have assumed that group members are randomly linked, such that the number of possible interactants for each player is $N - 1$. However, most networks are structurally differentiated. For a given population size, a high-density network in which each position exchanges with every other will be less conducive to collective action than a chainlike structure in which each has only two interactors who are not linked with each other. This chainlike pattern may be found on a street where residents relate primarily to their next-door neighbors or in an organization with a vertical chain of command or division of labor that imposes hierarchical and functional restrictions on interaction and access. Low density may also occur in a highly centralized structure where peripheral positions interact only with the core and not with each other. These structures

[11] Numerous behavioral experiments have indicated that reward is generally more efficient than punishment (e.g., Skinner 1953), but that is unrelated to the point here, which is entirely a consequence of the logic of the Prisoner's Dilemma and has nothing to do with the psychology of the actors.

increase strategic players' expectations of future engagement with current partners. They likewise reduce the number of choices that must be stochastically coordinated for adaptive actors to escape anomic equilibrium.

The advantage of structural differentiation is lost, however, if incumbents move easily among clusters of linked positions. The number of partners of each player thus depends on three factors: the size of the group, network density, and the degree of mobility. Consider the familiar rational choice argument that upward mobility provides an "exit," an alluring alternative to individual efforts on behalf of a disadvantaged collectivity (Blau and Duncan 1967, p. 440). More generally, Hirschman (1970) shows how exits undercut "voice" by giving rational actors a private solution to collective problems, one that does not require them to depend on the cooperation of others. For example, the availability of private schools undercuts the willingness of concerned parents to fight for better public schools (Hirschman 1970, pp. 45–46). This logic applies to social exchange as well. Those who expect to move off the block have less incentive to act like good neighbors.

Social learning theory suggests a nonrational explanation for the logic of mobility: exits tend to disrupt social ties. Social and territorial mobility may undermine collective action not only because of the lure of "getting ahead" but also because of the disruption of the stable relationships required to escape noncooperative equilibrium by stochastic collusion. Hence, the effects of mobility on a given group are more likely to be associated with "inflow" (from other classes or neighborhoods) than "outflow," as observed by Goldthorpe (1980, pp. 62–63) in his study of "working-class maturation."

It is a truism among sociologists that collective action is more likely to succeed among actors who are concentrated in a bounded social space, with common economic interests and multiple and overlapping ties, like the public employees in the imaginary "Centauri" proposed by Marwell et al. (1988, p. 505). Their simulations of contributions to public goods show how high interdependence increases the benefits of collective action, while physical and social proximity reduces the costs of communication and organization. The analysis of multilateral pairwise exchange suggests a different advantage. Compared to more fully rationalized structures, networks with concurrent ties entail more frequent exchanges with fewer people. Centaurians are kin to their friends, friends with their neighbors, neighbors with their workmates, and they all belong to the same church, PTA, and bowling league. Centaurians may grow tired of one another, but they enjoy the benefits of low coordination complexity.

They may also benefit from their lack of anonymity. The ability to differentiate interactants is a precondition for tacit collusion among rational actors, and Centaurians cannot easily disappear into the crowd. Simulations of adaptive behavior show that this precondition for tacit collusion is sufficient; if the players are able to differentiate among partners, rationality becomes superfluous. Suppose the actors are backward-looking and reactive yet able to recognize former partners, remember the choices and outcomes of their exchange, and careful not to generalize the lessons of that experience to their relationship with others. In other words, each player differentiates and modifies their cooperative propensities according to the identity of their interactant. Lock-in will now be no more difficult in N-way interactions than in the isolated two-player relationship simulated in figure 3. The size of the group will compromise stochastic collusion only insofar as it exhausts the capacity of the contestants to recognize former partners. In very low density networks, this differentiation is relatively easy, and the collective action problem is equivalent to a series of simultaneous two-player games. With maximum reinforcements, stochastic collusion will then occur within three iterations, no matter how large the network.

D. Evolving Structure: Exit, Loyalty, and Inequality

This finding has important implications for rational choice solutions to social traps. Rational choice theorists may have mistaken the ability to anticipate long-term consequences for what really matters in collective action, *the ability to tell people apart*. The mistake is easy to make since the former is useless without the latter. Tacit collusion requires the capacity to differentiate interactants, but we now see that if the actors have this capacity, they do not need the shadow of the future. Even stolidly myopic pragmatists can then be expected to find their way over the threshold and into lock-in. Collusion that succeeds on the basis of enlightened self-interest is likely to have succeeded without it.

Unfortunately, as the number of interactants increases, the actors find it increasingly difficult to keep careful track of who can and cannot be trusted. Moreover, those who fail to exercise foresight may also fail to see the need to file this information for future reference. In large and highly mobile groups, successful collusion may therefore depend on loyalty, the willingness to forgo opportunities to exit.

In rational choice theory, loyalty creates a vested interest in cooperating with the interactant (Hirschman 1970). Social learning theory suggests a different advantage. Adaptive players can facilitate stochastic coordination if they stop playing the field. Suppose the actors choose not only whether to cooperate or defect but also whether to continue the relationship or look elsewhere, and that their preference for partners, like their propensity to cooperate, is conditioned by the outcomes of the interactions. Neighbors do not interact equally with everyone on the street. While they may prefer the convenience of interacting with those next door or with those who share common attributes, it is entirely plausible that, all things being equal, they will prefer those with whom they enjoy a successful exchange. This includes of course those they are able to exploit, but now we assume that their victim, instead of retaliating in the next round, may simply walk away.

The original model can thus be elaborated by having the players modify their loyalty to their current interactor based on the same algorithm for transitional probabilities that modifies their propensities to cooperate or defect. If two randomly paired players happen to cooperate, the reward strengthens not only their willingness to cooperate again but also their propensity to stay with their current partner instead of searching for someone else. The players thus have a second payoff-conditioned propensity that shapes the structure of the network. The shape of the network in turn affects the feasibility of stochastic collusion.

It should not be surprising that relationships based on mutual attraction would be more conducive to solidarity than are random pairings. Rational actors can be expected to cooperate more readily with trusted associates than with strangers whose strategies are unknown and whom they may never see again. Learning theory arrives at a similar conclusion but for a different reason: any nonrandom pairings (and not just those produced by player preferences) reduce the number of decisions that must be stochastically synchronized to induce a mutually reinforcing collaboration and break free of anomic equilibrium. If everyone interacts randomly, anomie tends to prevail.

The simulations confirm this but also reveal an unexpected difficulty. When anomie prevails, everyone tends to interact randomly. The obstacle to cooperation is the instability associated with initiating stochastic collusion. Since players at equilibrium are noncooperative, escape typically begins with a series of *DD* rather than *CC*. While double defection has the same effect on cooperative propensities as *CC*, it has the opposite effect on mutual attraction. A sequence of double defections encourages both players to question the wisdom and propriety of their antisocial behavior, as well as the continuation of their relationship. The players may become more cooperative, but not with each other, thus undermining the synchronization needed to achieve lock in. Only attractions built on mutual cooperation will lead to stable interaction. Hence, players at noncooperative equilibrium may be expected to have fleeting relationships based on weak and fickle preferences and an aversion to suitors. The unstable pairings in turn impede escape from noncooperative equilibrium. In short, elective relationships appear to be more effective for preserving collusion where solidarity is already thriving than for getting collective action started in an anomic environment.

But now suppose the contestants cannot all afford to be equally choosy about whom they want to relate to. The payoffs in the Prisoner's Dilemma may be hypothesized to operate not only as reinforcements but also as *resources* that weight the preferences they condition. Better-endowed players enjoy the option of rebuffing those whom they find unattractive and waiting for someone more desirable. Resource-poor contestants, on the other hand, may not be able to afford that luxury. In social exchange, the more one has of what others want, the easier it is to find a client. Resources also confer the ability to choose and dismiss subordinates, as well as restrict or gain access to offices, neighborhoods, or clubs, as suggested in the large stratification literature on exclusionary practices and social closure. In short, the more resources players accumulate, the greater latitude they have in deciding with whom to share them. And depending upon what happens in the exchange, those resources may or may not continue to accumulate.

Forced interactions in which the victim cannot escape might be expected to promote exploitation compared to interactions where unwanted advances can be spurned. The paradox, however, is that it is the other way around. Figure 7 shows how asymmetric pairings are more conducive to collective action than are relationships based on mutual consent. The simulation begins with random preferences for partners and equal resource endowments. The players then modify their attraction toward their current interactor as play proceeds, just as they did in the mutual-consent game. As resource inequalities emerge, based on accumulated winnings and losses, the better-endowed players pick first and the weakest pick last. Figure 7 indicates the mean propensity of the eight players who begin the game at noncooperative equilibrium. Compared to a game with identical reinforcements ($\sigma = .1$ and $\gamma = 2$) but consensual pairings, all eight players lock in mutual cooperation in about half as many iterations. (Lock-in is achieved at $i = 43$, $i = 52$, $i = 90$, $i = 122$, $i = 158$, $i = 208$, and $i = 251$. Although the iterations required for lock-in remain unrealistically high, they rapidly decline as the magnitude of the reinforcement increases).

These simulations show once again how exits can undercut collective action, not for the reason given by rational choice theory, but because exits destabilize the emergent structure of the network. Consensual pairings cause relationships to be highly unstable at noncooperative equilibrium. This is not the case where exploitation is rewarded with the power to pick one's victim. The dominant player learns to hold on to those who typically cooperate and to abandon those who do not. Hence its victim will be relatively compliant initially, increasing the predator's accumulated winnings (thus enabling that player to remain dominant) as well as increasing the attraction to the victim. However, the victim eventually

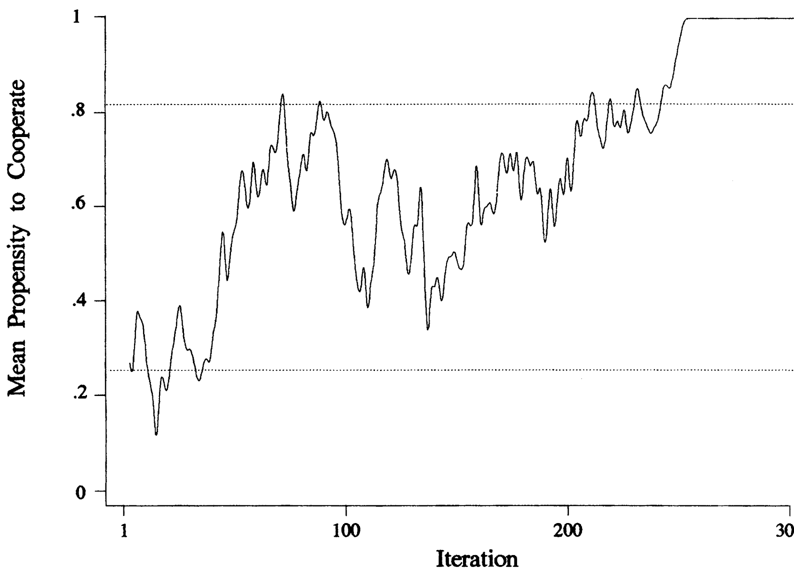

FIG. 7.—Effect of resource-weighted partner preferences on stochastic collusion ($\sigma = .1$, $\gamma = 2$, $N = 8$).

comes to prefer making life miserable for its tormenter, even though it must suffer as well. Before the predator goes on the prowl for a fresh patsy, there is the chance that both players will learn that more can be gained by collaboration. Pairings based on resource inequality are thus better able to withstand the destabilizing effects of the double defections needed to initially break out of anomic equilibrium, keeping the contestants together until stochastic collusion can begin to take hold. The simulations show that resourceful predators are more readily sobered when their victims cannot flee and must learn instead to fight, a compelling instance of the behavioral logic of exit and voice.

E. Invasion by Free Riders

Escaping the social trap is only half the difficulty. Once collective action is flourishing, the problem is reframed: How might solidarity be sustained in the face of invasion by free riders? Adaptive actors will not allow the appearance of a predator to alter their extant relationships so long as they are able to identify the intruder and alter their behavior accordingly. The difficulty arises where the free riders are hard to spot. The contamination is then likely to spread until it reaches every player linked to the intruder. Suppose a small, tightly knit neighborhood has secured universal compliance with the norm of tidying up after one's dog. No one likes to comply, but everyone likes to live where rules like this are effective. Then an unsocialized newcomer moves in. Each time a neighbor discovers an unpleasant violation in their driveway, their own enthusiasm for good citizenship declines, which in turn only increases the number of soiled shoes. It is not long before other norms begin to lose their grip, leading to a downward spiral in community obligation. The maxim is well known: "One rotten apple will spoil the barrel."

Unfortunately, the converse does not hold. One small candle turns out not to dispel the darkness. A good citizen who moves into a neighborhood of privatists will soon learn to play by the rules of the street. Privatism is readily exported; solidarity is not.

The solution to "backsliding" in the two-person game suggests one way that an invasion by inconspicuous free riders might be contained. While pragmatic (hence, decisive) responses are better at inducing stochastic collusion, the routinization of symbiotic exchange as socially appropriate behavior appears to be more helpful in maintaining solidarity in the face of invasion. Normative behavior slows the decay of solidarity, allowing victims of the intruder to recover their cooperative propensities through interaction with their comrades before being exploited again. Invasion by a free rider then turns the game into a variant of "freeze tag," but with the complication that victims of the intruder must be repeatedly "tagged" by their teammates to overcome the effects of a single encounter with "It." Thus, as more players become frozen, it is easier to freeze the rest. As an added risk, a victim that is frozen too long or too often is transformed into a confederate of It and then freezes his unsuspecting saviors. Retagging by teammates only restores solidarity if the victims of It remain almost certain to cooperate on the next move. Otherwise, interaction with former comrades leads to another incidence of exploitation rather than the mutual cooperation needed to recover from the effects of the previous tag. Rather than recovering through interaction with former teammates, they will instead spread the invasion further. Hence, the strategy is not likely to be effective except in very mild games where each player has many interactants and where the predator is badly outnumbered by cooperative players.

This is evident in figure 8, a relatively mild game with very small reinforcements ($\gamma = .15$, $\sigma = .01$). This models the low learning rates associated with attachment to rote decision rules, thus permitting ample opportunity for victims to recover. Consequently, the collectivists are able to sustain collusion for nearly 400 iterations (equivalent to 40 iterations with $\sigma = .1$), despite random interactions with a lone predator.

Figure 8 also reveals the fundamental weakness of this strategy. Cooperative players can only postpone the inevitable unless a way can be

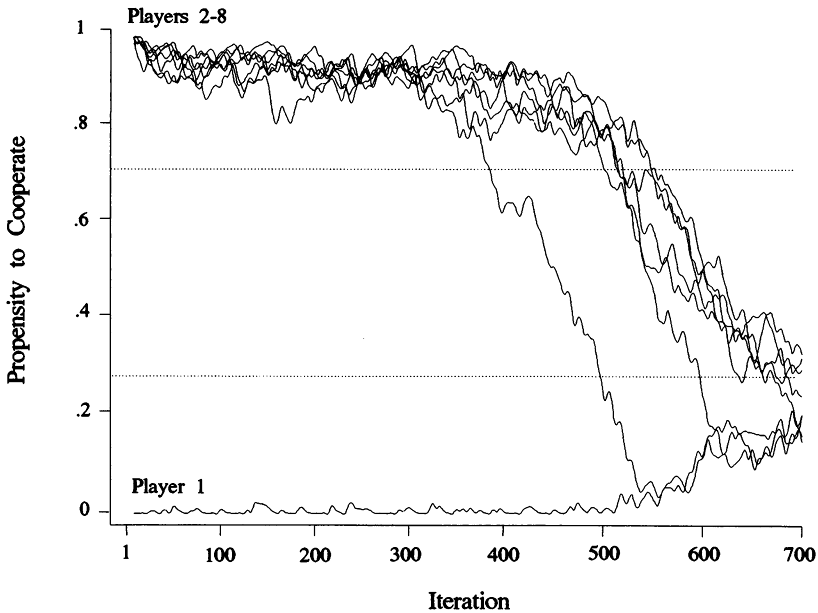

FIG. 8.—Collapse of normative solidarity caused by single defector (σ = .01, γ = 1.5, N = 8).

found to convert the intruder. Eventually, one of the victims becomes a confederate, after which solidarity quickly deteriorates. The problem is that the collectivists are overly nice. Their strong commitment to turning the other cheek helps contain the contamination but also encourages the predator. Hence it is only a matter of time before lock-in collapses. Normative solidarity appears to be more effective as an antidote for occasional backsliding than as a defense against predators.

IV. SUMMARY AND CONCLUSION

A simple two-person stochastic learning model shows how mutual cooperation might arise despite the risk that free riders may take advantage of altruistic gestures. The social trap lurking at the conjunction of micro- and macroprocesses turns out to be a noncooperative equilibrium that even badly myopic instrumentalists can escape if they are able to shift course decisively in response to immediate outcomes. No foresight is necessary; the behavioral model assumes only that their cooperative propensities are shaped over time by social sanctions and cues. The claim is not that higher-order reasoning is never effective in securing cooperation, only that it cannot always be assumed to occur. This approach is therefore applicable where actors do not grasp the strategic implications of an iterated Prisoner's Dilemma and must grope their way in response to social feedback. Social learning theory suggests the hypothesis that adaptive actors are led into a social trap more readily than are those able to look several moves ahead, but they are also better at finding their way out.

Once lock-in is attained, the problem is how to sustain collusion in the face of occasional backsliding. The simulations suggest that norms of civility are ineffective in getting cooperation started and are unlikely to emerge at noncooperative equilibrium due to the inconsistent sanctioning of behaviors. However, once the contestants have learned to collaborate as a pragmatic alternative to self-defeating conflict, the routinization of cooperation may promote the forgiveness needed to hold the alliance together over bumps in the road.

Extending the principle of stochastic collusion to multilateral encounters, computer simulations demonstrate the growing difficulty of securing collusion as the number of choices that must be synchronized increases. The findings corroborate the structural conditions usually thought to be conducive to collective action but show that the hypothesized effects of anonymity and "one-night stands" do not depend on the assumption that the actors appreciate the strategic implications of future exchange. The important cognitive skill is not the ability to estimate long-term cumulative returns but the ability to differentiate among multiple interactants. If the conditions required for tacit collusion in N-way interactions obtain, it is not necessary for self-interest to be enlightened by an appreciation of the long-term stakes.

If the players fail to differentiate, secondary sanctions (exogenous to the payoffs) may be needed to solve the coordination problem. Sanctions should not be focused on specific actors but should be broadly applied and timed to create a palpable nodal point for the synchronization of responses. Negative sanctions are needed where mutual defection is suboptimal but painless. With very low rates of cooperation, the reinforcement of civic virtue encourages free riders as well. Once cooperation gains a foothold, however, this risk declines, and additional positive reinforcement can help pull the players over the threshold and into lock-in.

The model also has implications for the problem of sustaining collective action in the face of incursion by unsocialized newcomers. Adaptive actors risk generalizing the lessons of betrayal, causing the contamination to quickly spread throughout the network. Resistance to invasion therefore requires a period of consolidation after lock-in is attained so that cooperation can be routinized through repeated reinforcement. The response to occasional betrayal may then be sufficiently attenuated so that victims of an intruder will be able to recover their solidarity through interaction with their comrades before being exploited again. Still, it is only a matter of time before they succumb.

These findings may be formalized as a series of hypotheses that might be tested using laboratory experiments. The purpose of the experiments is not to determine whether the subjects fail to exercise strategic foresight, but to test the conditions predicted to promote lock-in in the absence of opportunities for tacit collusion. Hence, it may be useful to alter the ability of the subjects to engineer a tacit collusion by varying the disclosure of the requisite information about the interdependence of choices and payoffs (the technique used by Rapoport and Chammah [1965]).

The following hypotheses are suggested by computer simulations of the two-player game with $T > R > 0 > P > S$.

HYPOTHESIS 1.—The probability of lock-in increases with the magnitude of the sanctions and decreases with the severity.

HYPOTHESIS 2.—If the payoffs are sufficiently large, the probability of lock-in will not increase with full disclosure of the choices and payoffs.

HYPOTHESIS 3.—*CD* increases the probability of *DD* at the next move and decreases the probability of *CC*.

HYPOTHESIS 4.—*DD* is not an absorbing state (*DD* eventually leads one or both players to cooperate).

HYPOTHESIS 5.—All else being equal, the greater the temptation to exploit (T), the lower the incidence of *CD* prior to lock-in and the lower the probability of lock-in.

HYPOTHESIS 6.—The longer lock-in is sustained prior to a disturbance, the shorter the time to subsequent recovery.

The following hypotheses are suggested by simulations of the N-player game.

HYPOTHESIS 7.—If the subjects are unable to differentiate among multiple interactants, (*a*) the probability of lock-in will decrease exponentially with the average number of interactants of each player, and (*b*) full disclosure of the choices and payoffs will not increase the probability of lock-in.

HYPOTHESIS 8.—A secondary sanction (exogenous to the payoffs) is more effective in facilitating lock-in when divided among all actors at each iteration than when concentrated on one actor and rotated at each iteration.

HYPOTHESIS 9.—The longer lock-in is sustained prior to invasion by an unsocialized newcomer, the longer it takes for lock-in to collapse.

HYPOTHESIS 10.—The probability of lock-in increases when subjects are allowed to choose their interactant by mutual consent.

HYPOTHESIS 11.—The probability of lock-in increases further when subjects are allowed to choose their interactant unilaterally (based on accumulated winnings).

These last two experiments with elective pairings signal a vast new research agenda, in which the structure of the network is no longer a given but a variable. Since contestants can usually choose not only whether to cooperate or defect but also whether to interact or withdraw, game theory can be used to explore how social structure might evolve in tandem with the collective action it makes possible. The N-player games provide an effective tool for mapping the rich complexity of structural arrangements that might emerge among interdependent actors. Contrary to the conventional wisdom, the analytic leverage of the game paradigm comes not from its reliance on rational choice theory but from the formalization of this web of mutually contingent relationships. Game theory would thus appear to be especially promising for sociology, the social science discipline that has been most reluctant to embrace it.

REFERENCES

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic.

Basu, Kaushik. 1977. "Information and Strategy in Iterated Prisoner's Dilemma." *Theory and Decision* 8 (3): 293–98.

Blau, Peter, and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York: Free Press.

Boring, Edwin G. 1950. *A History of Experimental Psychology*. New York: Appleton Century Crofts.

Bush, R. R., and F. Mosteller. 1955. *Stochastic Models for Learning*. New York: Wiley.

Carnevale, Peter J., Dean G. Pruitt, and Patricia I. Carrington. 1982. "Effects of Future Dependence, Liking, and Repeated Requests for Help on Helping Behavior." *Social Psychology Quarterly* 45 (1): 9–14.

Clark, M. L. 1983. "Effect of Confederate Sex and Status on Competitive Behavior of Male and Female College Students." *Journal of Psychology* 113 (2): 191–98.

Coleman, James S. 1986. "Social Structure and the Emergence of Norms among Rational Actors." Pp. 55–84 in *Paradoxical Effects of Social Behavior*, edited by A. Diekmann and P. Mitter. Heidelberg, Wien: Physica.

———. 1987. "Free Riders and Zealots." Pp. 59–82 in *Social Exchange Theory*, edited by Karen Cook. Newbury Park, Calif.: Sage.

Cross, John. 1983. *A Theory of Adaptive Economic Behavior*. Cambridge: Cambridge University Press.

Diekmann, Andreas. 1985. "Volunteer's Dilemma." *Journal of Conflict Resolution* 29:605–10.

Elster, Jon. 1979. *Ulysses and the Sirens: Studies in Rationality and Irrationality*. Cambridge: Cambridge University Press.

———. 1985. *Making Sense of Marx*. Cambridge: Cambridge University Press.

Flood, Merrill, Karl Lendenmann, and Anatol Rapoport. 1983. "2 × 2 Games Played by Rats: Different Delays of Reinforcement as Payoffs." *Behavioral Science* 28:65–78.

Fox, John, and Melvin Guyer. 1978. "Public Choice and Cooperation in N-Person Prisoner's Dilemma." *Journal of Conflict Resolution* 22 (3): 469–81.

Gardner, Rick M., Terry L. Corbin, and Janelle S. Beltramo. 1984. "The Prisoner's Dilemma Game and Cooperation in the Rat." *Psychological Reports* 55:687–96.

Goldthorpe, John. 1980. *Social Mobility and Class Structure in Modern Britain*. Oxford: Clarendon Press.

Hardin, Russell. 1982. *Collective Action*. Baltimore: Johns Hopkins University Press.

Hechter, Michael. 1987. *Principles of Group Solidarity*. Berkeley and Los Angeles: University of California Press.

Heckathorn, Douglas D. 1988. "Collective Sanctions and the Creation of Prisoner's Dilemma Norms." *American Journal of Sociology* 94:535–62.

———. 1989. "Collective Action and the Second-Order Free-Rider Problem." *Rationality and Society* 1 (1): 78–100.

Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty*. Cambridge, Mass.: Harvard University Press.

Homans, George C. 1961. *Social Behavior: Its Elementary Forms*. New York: Harcourt Brace Jovanovich.

Lacy, William B. 1978. "Assumptions of Human Nature, and Initial Expectations and Behavior as Mediators of Sex Effects in Prisoner's Dilemma Research." *Journal of Conflict Resolution* 22 (2): 269–81.

Lindskold, Svenn, and Marta L. Finch. 1981. "Styles of Announcing Conciliation." *Journal of Conflict Resolution* 25 (1): 145–55.

Luce, R. Duncan, and Howard Raiffa. 1957. *Games and Decisions*. New York: Wiley.

Macy, Michael W. 1989. "Walking Out of Social Traps: A Stochastic Learning Model for Prisoner's Dilemma." *Rationality and Society* 1:197–219.

———. 1990*a*. "Reply to Swistak." *Rationality and Society* 2 (2):242–47.

———. 1990*b*. "Learning Theory and the Logic of Critical Mass." *American Sociological Review* 55 (December): 809–26.

Marwell, Gerald, Pamela E. Oliver, and Ralph Prahl. 1988. "Social Networks and Collective Action: A Theory of Critical Mass. III." *American Journal of Sociology* 94 (3): 502–34.

Maynard-Smith, John. 1982. *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

Murnighan, J. Keith, and Alvin E. Roth. 1983. "Expecting Continued Play in Prisoner's Dilemma Games: A Test of Several Models." *Journal of Conflict Resolution* 27:279–300.

Oliver, Pamela. 1980. "Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations." *American Journal of Sociology* 86 (6): 1356–75.

———. 1984. "Rewards and Punishments as Selective Incentives: An Apex Game." *Journal of Conflict Resolution* 28 (1): 123–48.

Oliver, Pamela E., and Gerald Marwell. 1988. "The Paradox of Group Size in Collective Action: A Theory of Critical Mass. II." *American Sociological Review* 53(1): 1–8.

Olson, Mancur. 1965. *The Logic of Collective Action*. Cambridge, Mass.: Harvard University Press.

Opp, Karl-Dieter, 1986. "The Evolution of a Prisoner's Dilemma in the Market." Pp. 149–68 in *Paradoxical Effects of Social Behavior*, edited by A. Diekmann and P. Mitter. Heidelberg, Wien: Physica.

Parsons, Talcott. 1968. *The Structure of Social Action*. New York: Free Press.

Pincus, Jeffrey, and V. Edwin Bixenstine. 1977. "Cooperation in the Decomposed Prisoner's Dilemma Game: A Question of Revealing or Concealing Information." *Journal of Conflict Resolution* 21 (3):519–30.

———. 1979. "Cognitive Factors and Cooperation in the Prisoner's Dilemma Game." *Psychological Record* 29 (4): 463–71.

Platt, I. 1973. "Social Traps." *American Psychologist* 28:641–51.

Rapoport, Anatol, and Albert M. Chammah. 1965. *Prisoner's Dilemma: A Study in Conflict and Cooperation.* Ann Arbor: University of Michigan Press.

Raub, Werner, and Thomas Voss. 1986. "Conditions for Cooperation in Problematic Social Situations." Pp. 85–104 in *Paradoxical Effects of Social Behavior,* edited by A. Diekmann and P. Mitter. Heidelberg, Wien: Physica.

Scott, John Finley. 1971. *Internalization of Norms: A Sociological Theory of Moral Commitment.* Englewood Cliffs, N.J.: Prentice-Hall.

Selten, Reinhard. 1975. "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games." *International Journal of Game Theory* 4:25–55.

Skinner, B. F. 1953. *Science and Human Behavior.* New York: Macmillan.

Steinfatt, Thomas M. 1973. "The Prisoner's Dilemma and a Creative Alternative Game: The Effects of Communications under Conditions of Real Reward." *Simulation and Games* 4 (4): 389–409.

Swistak, Piotr. 1990. Comment on Macy. *Rationality and Society* 2 (2): 238–41.

Tognoli, Jerome. 1975. "Reciprocation of Generosity and Knowledge of Game Termination in the Decomposed Prisoner's Dilemma Game." *European Journal of Social Psychology* 5 (3): 297–313.

Turner, Jonathan. 1987. "Social Exchange Theory: Future Directions." Pp. 223–38 in *Social Exchange Theory,* edited by Karen Cook. Newbury Park, Calif.: Sage.

Winter, Sidney G. 1986. "Comments on Arrow and on Lucas." Pp. 243–50 in *Rational Choice: The Contrast between Economics and Psychology,* edited by Robin M. Hogarth and Melvin W. Reder. Chicago: University of Chicago Press.

Zagare, Frank C. 1984. *Game Theory: Concepts and Applications.* Beverly Hills, Calif.: Sage.