The use of game-theoretical explanations and predictions in empirical social science research is often thought to be precluded by the unrealism of the cognitive and computational capabilities that mathematical game theory imputes to its idealized "players" as well as by the prohibitive information costs that an attempt to reconstruct these cognitions and computations would impose on researchers. The article tries to show that these misgivings are exaggerated. Under realistic conditions actors will often be able to pragmatically approximate complete-information conditions regarding each other's strategy options and payoffs. Moreover, empirical research will, in many situations, be able to reconstruct actors' relevant opportunities, perceptions and preferences from socially constructed institutions, norms, and expectations that have always been the subject of mainstream social science research.

Games Real Actors Could Play

THE PROBLEM OF MUTUAL PREDICTABILITY

FRITZ W. SCHARPF
Max Planck Institut für Gesellschaftsforschung, West Germany

Game theory has become a powerful tool for the normative analysis of interdependent choices. It is also clear that game models representing the reflexive rationality of actors anticipating each others' choices correspond better with many of the real-world situations studied by social scientists than is true either of single-actor decision theory or of neoclassical economics with its focus on the aggregate effects of atomistic competition among large numbers of actors. Thus, if the rational choice approach should find a permanent place in the social sciences, it is more likely to succeed through game-theoretical applications than through the extension of neoclassical economics to new domains.

But while mathematical game theorists are generally content with asserting the normative validity of their models, it is far from clear what social scientists interested in the explanation and prediction of real-world phenomena would have to gain from analyses specifying how actors with perfect

471

information and unlimited cognitive capabilities *ought* to maximize arbitrarily defined utility functions. Thus, in order to assert the empirical usefulness of game theory, one would need to have satisfactory answers to three interrelated sets of problems:

1. How could real-world actors obtain the information that game-theoretic models presuppose they have?
2. How could real-world actors perform the calculations that game-theoretic solutions require?
3. How could researchers reproduce the information available to, and the calculations employed by, real-world actors?

These are difficult problems. Yet, contrary to a widespread prejudice among empirically oriented social scientists as well as among game theorists (Binmore 1987; Ryll 1989), it is by no means a foregone conclusion that adequate solutions cannot be found, if not generally, then at least for many kinds of situations. And given the potential explanatory power of game models, it seems highly worthwhile to explore the conditions under which one might have good reasons to use them in social science research.

BOUNDED RATIONALITY AND INCOMPLETE INFORMATION

Classical game theory assumes that there are no limits to players' computational capabilities, and that the strategy options and payoffs of every player should be "common knowledge" (Aumann 1976) among all players. To conform to these assumptions, real actors would need to know each other's action resources and constraints in order to arrive at a mutual understanding of their respective sets of available strategies. They would further have to estimate the causal consequences of all n-tuples of strategies in order to determine the physical outcomes obtainable. And they would need to know each other's perceptions and subjective valuations of these outcomes in order to construct the "payoff matrices" which could finally be subjected to the mathematical algorithms of game-theoretical analysis. For real actors, operating under the dual constraints of constrained or costly information and of bounded rationality, these are demanding requirements indeed, whose lack of realism is—perhaps too readily—granted by leading game theorists:

Classical game theory cannot handle games with incomplete information at all. This obviously poses a very serious limitation since virtually all real-life game situations involve incomplete information. In particular, it very rarely happens

that the participants of any real-life social situation have full information about
each other's payoff functions. Uncertainty about the strategies available to the
other players is also quite common. (Harsanyi and Selten 1988, 10)

In order to reduce the demands on available ex-ante information, John
Harsanyi (1967/68) introduced a "Bayesian" solution concept that merely
requires that players should be able to start with some subjective probability
estimates of each others' strategies and payoffs. By this substitution of "risk"
for "uncertainty," games with "incomplete information" will be transformed
into games with "imperfect information"[1] that can be analytically handled
by classical game theory (Selten 1982; Rasmusen 1989). But while the
analytical validity of this ingenious solution is generally accepted, its appli-
cability to "real-life social situations" remains very much in doubt. It replaces
allegedly unrealistic assumptions about actors' access to information with
even greater demands on their computational capacity, since they are ex-
pected to generate a substitute for complete information by tracing, analyz-
ing, and sequentially updating their understanding of mutually interdepen-
dent subjective probabilities (Eichhorn 1982; Colman 1982, 23-24; Friedman
1986, 116). Thus, while relaxing the information constraints, this solution
concept pushes even harder against the constraints of bounded rationality, so
that it seems fair to conclude that "Bayesian rationality has its place in
normative theory but a more realistic approach should be developed for
descriptive purposes" (Selten 1985, 82).

But in what respects should a "more realistic approach" differ from
ordinary game-theoretical assumptions and solution concepts? The difficul-
ties can all be traced to the fact that "rational-analytic" game models must
claim to be true representations of the mental operations of perfectly rational
actors. While the actors in neoclassical economics may be happily ignorant
at least of the analytical machinery of the "invisible hand," in game theory,
all information and all solution algorithms that are used by the analyst are
also imputed to the players themselves. Thus bounded rationality is a much more
serious problem for game-theoretical applications than it is in economics.

That may explain the recent interest in variants of game-theoretical
explanations that disconnect the analytical machinery from the calculations
of rational actors. An extreme case is evolutionary game theory as it has been
reimported from biology into the social sciences. Dispensing altogether with
the assumption of actor rationality, it interprets "strategies" either as geneti-
cally fixed behavioral propensities of the human species (Frank 1988) or as
"hardwired" routines and standard operating procedures of organizations
(Hannan and Freeman 1977, 1984). Evolutionary selection is then relied on

to produce outcomes at the level of populations of strategies that are identical with the equilibria identified by rational-analytic game theory (Maynard Smith 1982; Axelrod 1984). It should be clear, however, that the success story of evolutionary game theory in biology can be emulated by the social sciences only in areas where it can be shown that a rigorous selection mechanism exists that will, in fact, weed out behavioral traits or strategies of inferior "genetic fitness." While such conditions may be approximated in the competition among firms in near-perfect markets (Alchian 1950; Friedman 1953), we surely have no reason to simply assume their universal presence in social interactions (Elster 1983).

The problem is mitigated in some variants of the evolutionary approach that rely on "game learning" as a selection mechanism. By assuming that actors, while not capable of ex ante optimization, have at least the intellectual ability of distinguish ex post between more and less favorable outcomes, they are not forced to rely entirely on external selection for the retention and diffusion of successful solutions (Schotter 1981; Nelson and Winter 1982; Witt 1986, 1989; Macy 1989). Thus local optima can be reached through trial and error (Hernes 1976) or "gradient climbing" (Elster 1979), and "lock-ins" (David 1985) or "joint-decision traps" (Scharpf 1988, 1989) may be explained as stable equilibria even if actors could not calculate their optimal strategies in advance. Given the existence of multiple equilibria or local optima, however, outcomes will be historically contingent or "path dependent," and there is no assurance that the equilibrium actually obtained should be Pareto-efficient. Moreover, game-learning models presuppose long iterations of identical encounters under relatively stable environmental conditions, since stable equilibria may never be reached if the speed of trial-and-error learning is slow in comparison to the rate of social change (Elster 1983, 64). Thus, while evolutionary and game-learning models may be pertinent for certain types of real-world social interactions, there is still a need for analytical models that make greater (but not unlimited) demands on the specifically human capability for anticipation and purposeful action in unique or rapidly changing situations.

If the constraints of bounded rationality are to be respected, however, such models must disable those elements of the analytical machinery that cannot plausibly be imputed to real-world players. Thus I am convinced that empirical applications should not resort to Bayesian solutions in the face of incomplete information. Instead, they must still rely on the computationally less demanding solution concepts of "classical game theory" even though these have been superseded in mathematical game theory by Harsanyi's innovation.[2] But that presupposes that there should be other ways, accessible

to real-world actors, of coping with the endemic problems of incomplete information.

A THOUGHT EXPERIMENT

In order to show that pragmatically satisfactory solutions to the complete-information problem are indeed available, I will begin with a counterfactual thought experiment in which it is assumed that actors know the objective state of the world and their own minds but are ignorant of what goes on in the minds of other actors. Thus actors are assumed to be aware of the objectively feasible strategy options of all players and of the physical outcomes associated with all strategy n-tuples, and they are able to derive their own payoffs from these outcomes. But players are not assumed to have any information about the subjective perceptions and preferences of other players and hence about their effective payoffs. Yet, from a strategic point of view, this is precisely the most important aspect of uncertainty.

How should rational actors respond to this type of ignorance? It is a basic fact of the human condition that actors are vulnerable to malevolent acts (including deception, slander, theft, arson, robbery, and murder) that are objectively feasible for practically anyone whom they encounter. Thus the only thing that stands between their present endowment and extreme losses is the fact that most people whom they encounter do not have a preference for hurting others. Being ignorant of these preferences, however, even risk-neutral actors would have reason to proceed with extreme caution. More specifically, being ignorant of others' payoffs, rational players would have to base their own choice of strategies entirely on the analysis of their own payoff matrix.[3] If that should reveal a possibility of severe losses,[4] they would do well to eliminate all strategies from their own repertoires which do not meet the *maximin*, or perhaps the *maximin regret* criterion (Colman 1982, 22-30; Hamburger 1979, chap. 3). In other words, their responses would be dictated by worst-case scenarios and by the "capabilities-not-intentions" logic of "rational generalship" (Colman 1982, 51).

For our purposes, it is important that these rules of generalized caution would generate highly predictable outcomes under conditions of incomplete information, even in single-shot encounters. But how would such outcomes differ from solutions of the same game played under conditions of complete information? It is easy to see that some opportunities for optimization must be lost if players cannot anticipate each other's moves. Yet whether these would have been opportunities for cooperation or for exploitation seems to

depend entirely on the character of the game that is being played. Take the Chicken game, for instance, where genuine uncertainty about the payoffs of the other side would lead either player to adopt a maximin strategy that would, in effect, assure a cooperative outcome. Or take Reinhard Selten's (1978) famous chain-store paradox where each one of 20 small, independent stores must separately decide whether to challenge a monopolist by entering its market, while the monopolist must then choose between acquiescence and a costly price war that will destroy the challenger. While the monopolist would acquiesce under complete-information conditions, any uncertainty about its payoffs would deter potential challengers from exploiting the monopolist's dilemma (Kreps and Wilson 1982).

By the same token, however, players are also unable to exploit opportunities for cooperation arising from an undiscovered harmony of their interests. Even under the benign conditions of the Assurance game, uncertainty about the payoffs of the other side would render defection the preferred choice. The same is true in Battle of the Sexes and, a fortiori, in Prisoner's Dilemma, where cooperation could at best arise from the correct perception of the other side's long-run self-interest when the game is indefinitely iterated (Axelrod 1984). There is also a good deal of experimental evidence showing that the likelihood of cooperation is reduced when players are kept ignorant of each others' payoffs (Rapoport and Chammah 1965; Guyer and Rapoport 1969; Hamburger 1979, 238-42). The implication is that games can indeed be played with incomplete information even if players are unable to resort to Bayesian solutions—but that their likely outcomes will be individually "unprofitable" (in the sense that players cannot expect to receive more than their maximin payoffs; Harsanyi 1977, 116, 138) and often socially inefficient as well.

But while generalized caution may be descriptive of interactions among superpowers during the cold-war period (Snyder 1971; Jervis 1978), or of encounters in big-city streets after dark and in 18th-century Naples (Pagden 1988), our thought experiment should not be mistaken as a generally applicable model. Quite evidently, maximin strategies are not used in most real-world encounters, and profitable games are, in fact, often played. However, by providing a baseline model of interactions under complete payoff uncertainty, our thought experiment suggests that profitable interactions cannot be taken for granted but need to be explained. It also suggests that an adequate explanation would have to identify mechanisms that are able to reduce uncertainty by increasing (a) the ability of actors to predict which strategies other players will, in fact, consider and (b) the ability to predict the payoff functions guiding other players' choices among the strategies consid-

ered. Such predictability-increasing mechanisms could be either endogenous or exogenous to game-theoretical models. In the first case, they would have to be reconstructed from the rational choices of players at the microlevel; in the second case, the search would be extended to influences of the preexisting societal and institutional environment in which players and their games are embedded.

THE INDIVIDUALISTIC
CONSTRUCTION OF PREDICTABILITY

Even though most real-world interactions are, in fact, institutionally embedded, individualistic constructs may throw light on the evolution of norms and institutions (Ullmann-Margalit 1977; Kliemt 1986; Shepsle 1989), and they may help to explain international and other types of "lawless" interactions. And even when institutions are in place, complex, dynamic, and pluralistic modern societies could not exist without wide margins of individual choice and discretionary action. The question is how actors on their own could develop the converging expectations that are required for profitable interactions.

There is, of course, an easy answer to the problem of incomplete information: If players have an interest in coordinating their expectations, they should agree on their permissible strategies and communicate their own payoffs to each other—in other words, they should play a cooperative game. But then, how could they trust each other? If truthful communication and binding agreements were dependent on exogenous enforcement, we would have left the realm of individualistic construction—unless one could show how such enforcement mechanisms themselves could arise in interactions that do not depend on their prior existence. In game theory, that condition is expressed by the maxim that "any equilibrium, cooperatively reached or not, must be a noncooperative equilibrium" (Ordeshook 1986, 303).

This seems a reasonable requirement in a world without institutions. But solutions are made unnecessarily difficult if "noncooperative" games are defined by the absence of both communication and exogenous enforcement of agreements, as is the usual convention in game-theoretical literature. Instead, I will here follow Harsanyi in basing the distinction on the availability of exogenous enforcement alone, so that the definition of "noncooperative" solutions will not rule out communication and nonbinding commitments among players. I will also follow Harsanyi's (1977, 115) further suggestion that every game may be preceded by a (possibly tacit) bargaining

game in which players will try to achieve agreed-on definitions of their allowable strategies and relevant payoffs—just as parlor games are often preceded by discussion and agreement on the applicable rules (Shepsle 1989).

By itself, of course, that would not yet overcome the crucial problem of how players could be able to trust each other's communications and commitments in the absence of exogenous enforcement, nor does it help much to know that some agreements may be self-enforcing (Harsanyi 1977, 110). While in certain classes of games (e.g., Assurance) agreements reached in preplay bargaining would define a strong equilibrium from which no one has an interest to depart, in other games (e.g., Chicken) players would profit from misrepresenting their preferences and/or from defaulting on their commitments.[5] However, given their initial ignorance of each other's payoffs, players could not know which type of game they are, in fact, playing. Thus, since they could not distinguish trustworthy communication from deception, they would remain locked in mutual distrust in both types of situations.

In short, players cannot, on their own, create the conditions of mutual credibility when the individual game is considered in isolation. But they may be able to do so if the game is placed within a wider context. Among game theorists, it is a well-known "folk theorem" that infinite[6] iteration may change the outcome of otherwise dilemmatic games because in the resulting "super-game," players may make their next move contingent on their opponents' previous choice. Thus it has been shown for the iterated Prisoner's Dilemma that cooperation may emerge as a self-enforcing equilibrium if players follow a "tit-for-tat" strategy that responds in kind to each instance of the opponents' cooperation or defection (Axelrod 1984; Taylor 1987). The same logic, it is suggested, may also help to solve the problem of truthful communications and credible commitments under conditions of incomplete information.

However, since not all real-life games have the form of Prisoner's Dilemma, and since infinite or even indefinite iterations of identical games are rare in any case, the logic must be modified to become more generally applicable. The starting point is Harsanyi's distinction between the primary game that is, in fact, being played and the preplay bargaining game in which payoffs may be communicated and strategies agreed on. If that is granted, it is plausible to assume that many primary games, even though quite different from one another, may still be embedded in a series of continuing interactions among the same players. Under such conditions, all single instances of preplay bargaining may be conceptualized as belonging to a sequence of more abstract (and hence more uniform) "truth games"[7] that are, in fact, indefinitely iterated. In these, players must choose whether they will tell the

	C	D
C	3 3	4 2
D	2 4	1 1

CHICKEN

	C	D
C	3 3	4 1
D	* 2 4	1 2

/ PD

	C	D
C		2
D	4	

	C	D
C	4 3	3 1
D	1 4	2 2

PD / ASSURANCE

	C	D
C	* 4 4	3 1
D	1 3	2 2

	C	D
C		
D	4 1	

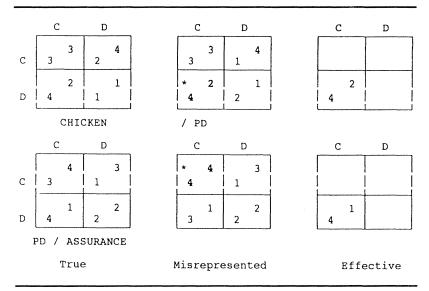True   Misrepresented   Effective

Figure 1:   True and misrepresented payoffs and effective outcomes.
NOTE: * = outcomes expected under the assumed payoffs.

truth and keep commitments and whether they will trust others' communications and commitments.

Now, if all instances of default could be detected immediately after the fact, each player would be able to respond in the next round to others' previous performance in the truth game. If we could further assume that players do, in fact, have a common interest in being able to play profitable games based on trustworthy communications and commitments, the game-theoretical folk theorem suggests that they should be able to reach that outcome through individually rational strategies (Fudenberg and Maskin 1986). But that conclusion is based on two critical assumptions that need to be discussed. First, the folk theorem will not help if others' defection cannot be detected before one's own next move—which has been a serious problem in arms-control negotiations. But even when the actual choice of *strategies* can be verified after the fact, there will often be no possibility of direct verification for the *payoffs* that have been communicated. While their truth may sometimes be indirectly confirmed or disconfirmed by the observed strategies, there are also cases where even that is not possible (see Figure 1).

Take the example of a Chicken game in which the "Row" player misrepresents one's own payoffs as being of the Prisoner's Dilemma variety. If

"Row" is believed, "Column" will expect "Row" to pick D as the dominant strategy and will choose C to avoid one's own worst-case outcome. As this will then allow "Row" to, in fact, play D in accordance with true (Chicken) preferences, the deception cannot be discovered.[8] A counterexample is provided by a game where "Column" has Assurance-type and "Row" PD-type preferences. If "Row" pretends to also have Assurance preferences, "Column" will be deceived into cooperation. But if "Row" then defects in accordance with true PD preferences, the deception will be revealed.

The examples can be generalized to similar types of games in which "reputation" makes a difference. While a false reputation for being cooperative will be destroyed if it is exploited, a false reputation for toughness can survive if the threat is believed and the bluff not called. In such cases, the deceived party will be induced to accept a poor bargain that is then kept by the other side. As far as I can see, there is no generally effective protection against the misrepresentation of payoffs in threat games—except, of course, the fallback strategy of generalized caution. But it is important to realize that the victim of a false threat is, in fact, no worse off than this person would have been by disbelieving all communications and following a maximin strategy to begin with.[9] So the real dangers that one must guard against are not false threats but false promises—which may, in fact, lead to outcomes that are below the maximin threshold. But these will show up after the individual play and thus can be sanctioned by appropriate reactive strategies in indefinite iterations of the truth game.

But what should this reactive strategy be? For iterations of the Prisoner's Dilemma, Michael Taylor (1987) and Robert Axelrod (1984) showed that a provokable but forgiving "tit-for-tat" strategy is superior not only to pure cooperation and pure defection but to an unforgiving response of "eternal damnation" after a single instance of defection. However, if the type of primary game that is being played is not initially known, but is drawn from a population that includes not only the PD but Assurance, Battle of the Sexes, Chicken and other games, tit-for-tat is not a plausible strategy for the truth game. In fact, a player who was known to be committed to tit-for-tat would become the perfect target for exploitation—deceived whenever the underlying game happened to be Chicken and willing to forgive after each play of Assurance (where telling the truth is in the other player's self-interest). By contrast, an unforgiving response ("once a liar, always a liar") to any single instance of deception would seem to be a superior strategy in the truth game.[10]

An opponent confronted with the unforgiving strategy would face the following choice: By being truthful and trustworthy, the player would maintain access to future games in which one's own communications are

trusted—and which, therefore, could be more self-profitable than encounters governed by generalized caution. A single case of (detected) deception, however, would thereafter restrict the player to interactions in which opponents would follow a strategy of generalized caution. Thus the temptation to deceive in an individual case would have to be weighed against the lesser profitability of all future interactions.[11]

By itself, however, the threat of an unforgiving strategy would not yet assure the truthfulness of communications and the trustworthiness of commitments in all situations. Punishment that must, in fact, be applied is often costly (Molm 1989). More specifically, the player who must, in fact, revert to generalized caution after a single instance of deception is also deprived of future profitable interactions even in games with common interests like Assurance. Thus the temptation to forgive in later games may be too great to maintain the deterrent effect. Moreover, even a credible threat to be unforgiving would not offer protection against a "confidence" player who would build up trust and then disappear after one big "killing."

Since breaches of trust do, in fact, occur, we have no reason to construct a theory that would exclude them altogether. Nevertheless, it seems worthwhile to identify the endogenous mechanisms that could explain why they do not occur much more often. Such mechanisms can be found by again extending the truth game to include an even wider context. In the real world, ongoing relationships are not restricted to dyadic interactions but often take place within a population of actors who are aware of, and able to communicate with, each other. Thus other players could also respond with an unforgiving strategy of generalized caution when it is their turn to play against a known defector. But since that form of sanctioning may also entail costs for these later players, what still needs to be explained is why they should be motivated to punish.[12]

One approach starts from the insight that the costs of punishment are maximized by the conventional assumption that players do not have the choice of refusing to play a game. Yet in many real-life situations, actors are, in fact, free to choose whether and with whom they are willing to interact (Tullock 1985). Thus, when confronted with a former defector, a player would have the following choices: (a) playing with the defector and trusting one's communications, (b) playing with the defector and resorting to generalized caution, (c) playing with someone else, and (d) not playing at all. If the latter two options are at all attractive, the cost of avoiding the first option could be much reduced, while the severity of punishment, and hence its likely deterrent effect on potential defectors, would, in fact, increase.

Under such conditions, the motive of self-protection may be all that is needed to assure the ostracism of players that are known to have defaulted on previous occasions (and are hence suspected of being less than totally trustworthy). It may even be sufficient to provide a commercial market for reliable information about the "credit worthiness" of other players whose past defaults one could not have observed directly (Milgrom, North, and Weingast 1988). And if the cumulative (and appropriately discounted) disadvantage of being excluded from future profitable games is greater than the present gain from cheating in a single game, rational players will have reason to maintain their reputation for keeping promises and telling the truth even in single-shot encounters.

To conclude this exercise in speculative exploration: Even in the absence of exogenous enforcement and of exogenous motivations to punish, truth-telling and promise-keeping is an equilibrium solution of the "truth game" in ongoing groups of self-interested actors who are able to observe each others' interactions. And given the abstraction from the underlying primary games and its own simplicity, the solution of the truth game itself would not overtax the computational capacity of ordinary players. Moreover, the informational requirements of discovering and communicating instances of deception and broken promises, while not trivial, are assisted by the human interest in gossip and scandal and by the competitive motive of the news media to cater to that interest. And finally, the equilibrium solution of the truth game also seems robust in the sense that even if not all cheaters will be caught, those that are caught are still likely to be punished.[13] In other words, the construct of a truth game is able to explain how self-interested and boundedly rational actors are able, under quite realistic conditions, to create an environment in which complete-information conditions are, in fact, approximated.[14] Thus we have reason to think that the solution concepts of "classical game theory" may indeed have considerable relevance for the analysis of real-life interactions.

But even if that is granted, preplay bargaining by itself cannot *generally* assure the empirical applicability of game-theoretical models. First, the postulated preconditions for the enlarged truth game, low-cost choice among partners and low-cost monitoring of the past performance of potential partners, will often not exist.[15] Second and more important, the assumptions introduced could only assure the *credibility* of communications and agreements, but they do nothing to reduce the *information costs* associated with idiosyncratic perceptions and preferences. To construct common-knowledge conditions on this foundation alone would impose enormous burdens on the communicative capacities of real-life actors who would need to recollect,

specify, transmit, and correctly interpret all decision premises that are poten-
tially relevant in characteristically ambiguous interaction situations. And
even if such feats of communication are conceivable among parties with a
long history of close association, they would create practically insuperable
problems for empirical research. Game-theoretical explanations might thus
be entirely valid in principle, and at the same time pragmatically useless, if
they should in every case depend on a precise reconstruction of the idiosyn-
cratic (and mostly tacit) understandings among specific actors.

THE SOCIAL CONSTRUCTION OF PREDICTABILITY

Thus we should search further for solutions to the common-knowledge
problem that will reduce the information costs for actors as well as for
researchers. In that search we must also cross the artificial dividing line
between rational choice and conventional social science approaches. In
trying to clarify the preconditions of social order, social scientists working
in the functionalist, institutionalist, or symbolic interactionist traditions have
always focused on social, cultural, economic, and political institutions, rules,
and conventions as devices for limiting the range of arbitrary human choices
and idiosyncratic worldviews. While these devices were ignored in institution-
free neoclassical models, they are now increasingly recognized by rational
choice theorists as essential informational mechanisms increasing the pre-
dictability and facilitating the coordination of human action (Schotter 1981,
109ff, 140ff; Heiner 1983). They do so by reducing the vast range of
objectively feasible strategy options to much narrower subsets that actors can
be expected to consider in specific interactions and by substituting socially
defined worldviews and "common sense" interpretations (Dupy 1989) for
the potentially unlimited variety of subjective interpretations of situations
and outcomes.

There is by now a considerable body of rational choice (mostly game-
theoretical) analyses showing how social institutions and conventions[16] could
"spontaneously" emerge from the unplanned interaction of self-interested
individuals (Hayek 1960; Ullmann-Margalit 1977; Hardin 1980; Schotter
1981; Kliemt and Schauenberg 1984; Shepsle 1989; Witt 1989; Sugden
1989). At the same time, another body of rational choice analyses, usually
with a normative orientation, focuses on institutional arrangements and
norms created by purposeful collective action (Buchanan and Tullock 1962;
Brennan and Buchanan 1985; Ostrom 1987; Hardin 1988; Ostrom 1989).
Clearly, these approaches do not contradict each other. Institutions may

emerge and change spontaneously and may also be created and reformed through purposeful action. Often, both types of processes will coexist and their interaction may be one of the most important and least understood phenomena in the evolution of modern societies.

What is important here, however, is the fact that the outcomes of both types of institution-building processes are likely to be historically contingent. For purposefully created institutions, shaped by the "tastes" of their creators, that is perhaps obvious. But it is also likely that the noncooperative games[17] generating "spontaneous order" will have multiple, and often very many, equilibria (Schotter 1981, 12-13). Which of them is selected will likely be influenced by the point of departure and by subsequent historical accidents. There is also no assurance that the equilibria so selected will be Pareto-optimal (Sugden 1989). They are path dependent, and history is not necessarily efficient (March and Olsen 1989, 54-56).

While institutions and conventions, even if path dependent, will solve many of the information problems of actors who are socialized and enculturated into the historically specific context of their interactions, they can provide only a second-best solution for the information problems of empirical research. Since their specific content cannot be deduced from a set of universal assumptions, institution-based explanations can only aspire to the status of "sometimes-true theories" (Coleman 1964, 516-19) that remain dependent on time-and-space specific information about socially constructed strategy options and payoffs.

In some fields of highly structured and highly public interaction, it may be sufficient to draw on legal and customary rules and publicly documented outcomes—which may explain the persuasiveness of game-theoretical explanations of coalition formation and roll-call voting in Congress and its committees (Shepsle 1979; Shepsle and Weingast 1981, 1987; Denzau, Riker, and Shepsle 1985). In other areas, however, institutionalized systems of rules will not completely determine choices but will still constrain the range of strategies that actors and observers need to take into consideration (Ostrom 1986; Burns and Flam 1987). Of particular importance for game-theoretical analyses are "boundary rules" defining which players have access to a particular arena, and "decision rules" restricting the range of outcomes that can be determined in a particular arena and the range of choices that a particular player may influence.

But even if institutionalized capabilities and constraints will define available "opportunities," actors' choices among feasible strategies will also be influenced by what Jon Elster (1989) labeled "desires"—that is, by actors' perceptions of the situation and their preferences among outcomes (pp. 13-21).

Again, however, these "desires" need not be empirically determined de novo in every case. In some situations, the relevant definitions of *self-interest* may be so basic that they may plausibly be imputed to most actors. While Adam Smith tried to get by with two universal types of utility, "social approval" and "physical well-being," Siegwart Lindenberg (1989, 190) found it neces-sary to add the "minimization of loss." For corporate actors, a similar short list might include organizational survival and autonomy. Whenever these fundamental interests are seriously at stake in an interaction, they may be all that is needed to explain actors' choices in what has been called "single-exit" or "high-cost situations" (Latsis 1972; Zintl 1989). In many other situations, however, basic self-interest by itself will be less directly useful as a predictor of perceptions and preferences. Even if it should still be the sole motivating force of human action, its pursuit would not only be externally constrained but internally shaped by socially constructed criteria of relevance and right-ness that have been well described as a "logic of appropriateness" (March and Olsen 1989, 23f).[18] Among these criteria, we may distinguish norms defined either in terms of universalistic justice, morality, or fairness (Etzioni 1988; March and Olsen 1989) or in terms of socially defined roles and collective identities to which actors must refer for their "self-interested" evaluation of outcomes. At least the latter kind of group-specific norms is often sanctioned by the highly effective threat of collective disapproval and ostracism. Of similar practical importance are functionally specific orienta-tions defined by the codes of relevance and excellence associated with differentiated societal subsystems (Luhmann 1986), with socially coherent professions, or with the development of large technical systems (Mayntz and Hughes 1988). Outsiders will have learned to take these specific criteria for granted, and insiders will have internalized them through professional social-ization supported by the incentives associated with professional reputation and ostracism.

To the extent that these socially constructed orientations are common knowledge (or "common sense") among actors, they will reduce the infor-mation problem of researchers as well. Even without direct access to actors' cognitive maps, it will often be possible to ascertain their expectations of each other's available strategies and payoffs from existing descriptive stud-ies, from public records and available documents, or through "expert inter-views" with insiders or knowledgeable observers. For attempts at rational choice explanations of real-world phenomena, that surely is an invaluable help. Yet there are two caveats.

First, since we are not dealing with universal laws but with historically contingent invariances, the empirical validity of institution-based regularities

must be reconfirmed for every new context of interactions. Thus expectations derived from studies of Swedish industrial relations in the 1960s never held for British industrial relations, and they are no longer true for Sweden in the 1980s (Scharpf 1987). Second and even more problematic for empirical research is the fact that institutions and conventions can at best support highly standardized expectations, while we know that there are, in fact, many interactions where the players do know a good deal more about each other. As Mark Granovetter (1985) emphasized, specific encounters are often embedded not only in a general context of institutional rules and social norms but in ongoing personal relations and networks of such relations. When such ongoing relations do exist, the reliability of actors' expectations, and their trust in each other's commitments, may be raised far above the level that would be reasonable among even well-socialized strangers. This has important implications for empirical research.

RELATION-SPECIFIC ORIENTATIONS

"Profitable" games, we have seen, can only be played if players' payoff functions are common knowledge. By reverse implication, that also restricts the range of motives which even cooperative players are able to pursue to a subset of which all can safely presume mutual awareness. Socially constructed orientations provide a general basis for such mutual awareness. But, as Simmel ([1917] 1984, 33-39) pointed out long ago, the very generality of such orientations systematically excludes those aspirations in which human beings tend to differ from one another. Even if we should not share Simmel's pessimism that this must always discriminate against "higher" aspirations,[19] it is clear that generalization must exclude wide ranges of concerns which are valuable to actors and in which they could benefit from informed and trustful interaction and exchange. That constraint is relaxed in networks of ongoing relations, in which more idiosyncratic concerns may indeed become the object of interdependent action.

Networks, in other words, reduce transaction costs and increase the potential gains from interaction. By the same token, however, they again increase the distance between the information available to real-world actors and the information that is readily accessible to the researcher. Thus, while the explanatory power of game-theoretical models increases with the increase of shared information among the actors, the actual use of such explanations in empirical research becomes more difficult as actors move

from generalized to relation-specific expectations regarding each other's strategy options and payoffs. But that is not yet the end of our difficulties.

Like standard economics, game theory usually assumes an "individualis-tic" orientation according to which actors are only concerned with their own utility. Logically, of course, that would not exclude altruistic or competitive orientations since "utility" may be defined to include any concern that the actor may have for others' welfare. Thus others' losses may count as one's own gain in war or in competitive sports, and guardians or professional helpers may be bound by professional ethics to include the welfare of their wards or clients in their own utility functions (Barber 1983). For the purposes of empirical (or experimental) research, however, such an encompassing definition of utility is less helpful than an alternate conceptualization which distinguishes between the outcomes directly accruing to an actor and the "interaction orientation"[20] through which this actor defines one's (positive, negative, or indifferent) attitude toward the outcomes accruing to others.

The usefulness of the distinction was first emphasized by social psychol-ogists reporting on game-theoretical experiments in which subjects seemed to respond irrationally to the payoffs specified by the experimenter. Their choices appeared entirely rational, however, when interpreted as responses to an "effective matrix" obtained by a subjective transformation of the objectively "given matrix" (Kelley and Thibaut 1978, 14-17; Kelley 1983).

Among these transformation rules, the ones most often encountered in experimental research may be described as "individualistic," "competitive," "solidaristic," "egalitarian," "altruistic," or "aggressive" interaction orienta-tions.[21] Empirically, these subjective interaction orientations may change dramatically, even if the underlying objective circumstances (the "given matrix") remain relatively stable. Examples that come to mind are not limited to the emotional volatility of personal relations but would also include instances of fundamental changes in the perception of political and military constellations—as brought about by Sadat's visit to Jerusalem or by the recent de-escalation of the big-power arms race (Kahn and Kramer 1990).

From a research-pragmatic point of view, the separate treatment of inter-action orientations seems to offer three advantages. First, it will often be relatively easy to identify the objectively "given matrix," while information about actors' subjective perceptions of "effective matrices" is likely to be more elusive. Second, it seems possible that empirical and experimental research will ultimately be able to identify a standard set of interaction orientations with well-defined empirical indicators that could be used as a checklist in research. Finally, there is reason to hope that it may be possible

to identify generalizable conditions and mechanisms that control the switches from one interaction orientation to another that seem to occur without prior changes in the given matrix.

These are attractive prospects for empirical research. Moreover, as I have tried to show elsewhere (Scharpf 1989), at least two of the most frequently encountered nonindividualistic orientations, "competition" and "solidarity," will radically simplify "given" game constellations by transforming complex mixed-motive games into "effective" matrices that represent analytically much simpler zero-sum or pure coordination games. Thus, by focusing our attention on interaction orientations, we are more likely to discover games that will not overtax the cognitive and computational capabilities of real actors than one should expect on the basis of the individualistic assumptions of neoclassical economics and mathematical game theory.

* * *

On the whole, therefore, this essay ends on an optimistic note: Game theory is a powerful tool for the disciplined theoretical analysis of interacting choices. While its analytical power is often seen to depend on unrealistic assumptions about actors' cognitive and computational capabilities, these requirements are less forbidding when we consider actors' opportunities to pragmatically approximate conditions under which they have complete information of each others' strategy options and payoffs. What I have tried to show is that it is indeed possible to identify theoretically plausible mechanisms facilitating the individual and social construction of higher degrees of mutual predictability. When and where such mechanisms can be found to operate, game-theoretical models should, indeed, become useful tools of explanation and prediction in empirical social science research.

NOTES

1. While "completeness" refers to information about payoffs and strategies, "perfectness" characterizes information available at any point in a game about prior moves of players (including chance moves attributed to "nature").

2. I also believe that empirical applications should do without the concept of "mixed strategies" in single-shot games. The notion that players should resort to a random mixture of their pure strategies is highly counterintuitive, or even meaningless, in most noniterated game situations (Neisser 1952; Aumann 1987, 15); it is, in fact, not used by subjects, even in long iterations of experimental games where its application would have been profitable (Colman

1982, 77). Of course, by ignoring mixed strategies, we increase the number of situations for which game theory will be unable to identify a unique, or any, equilibrium solution. But what we lose in analytical power, we gain in greater realism.

3. Players thus cannot apply the central game-theoretical "rationality principle" which would direct them to anticipate the optimal choices of other players in determining their own best response (Hamburger 1979, 44). Of Harsanyi's (1977, 116-18) two sets of rationality postulates, "payoff dominance" and "rational expectations," they are able to rely only on the former.

4. The rule of generalized caution is not based on a "framing effect" implying the systematic overestimation of losses (Kahneman and Tversky 1984) but on the assumption that in encounters among strangers, the objective magnitude of potential losses is likely to exceed the magnitude of potential gains.

5. The temptation to misrepresent one's payoffs exists whenever players' most preferred outcomes differ from one another, and it is irrelevant when the other player has a dominant strategy. The temptation to default on an agreement exists when the agreed-on (Pareto-efficient) solution is not a noncooperative equilibrium. Thus both types of temptation will exist in Chicken and neither of them in Assurance, while only the temptation to misrepresent exists in Battle of the Sexes and only the temptation to default in Prisoner's Dilemma.

6. In finite iterations, rational players would defect on the last play and hence on the second-to-last play, and so on. Whether the backward induction of this "endgame effect" can be avoided in finite iterations with incomplete information is in dispute. (Fudenberg and Maskin 1986; Güth, Leininger, and Stephan 1988).

7. The idea of a "truth game" was introduced by Brams (1985, 119-26) in the context of his analysis of the arms race among superpowers. What is different here are the assumptions of iteration and of the heterogeneity of the underlying primary games.

8. Of course, if both should misrepresent their preferences in this fashion, and if each should believe the other, the deception would become visible—and both would be surprised to find themselves with a cooperative outcome in the C/C cell.

9. In other words, the relevant standard of comparison is not complete information ("If only I had known") but acknowledged ignorance ("If only I had not thought I knew").

10. Alternatively, one might consider a strategy that begins with generalized caution in first encounters and incrementally increases the level of one's trust in opponents' communications with each subsequent instance of demonstrated truthfulness (in games where deception would have been advantageous). This accumulated reputation for trustworthiness could be destroyed by a single instance of detected deception but might be rebuilt again (perhaps under more difficult conditions).

11. The same logic does, of course, apply to situations where different games are simultaneously ongoing among the same parties (McGinnis 1986).

12. Axelrod (1986) demonstrated in computer simulations that a metanorm which makes not punishing a defector a punishable offense could generate a stable equilibrium of universal punishment. But the emergence of this metanorm is not itself explained as the outcome of rational choices. Similarly, Witt (1986) merely showed how the tendency to punish defectors might be stabilized after it has, somehow, become dominant in a population.

13. The likelihood may be increased by genetically "hardwired" emotional responses against liars and cheaters (Frank 1988).

14. Moreover, by assuring the possibility of trustworthy commitments, the truth game may create the even more demanding preconditions of games with "truly perfect information" (Thompson and Faith 1980, 1981).

15. Thus nation-states, government organizations, political parties, unions, or employer associations do not usually have the option of avoiding each other, and the same is true of individuals in organizations or families. But these captive players may have excellent opportunities for monitoring and sanctioning.

16. Where a distinction is important, the line might be drawn between "conventions" defining self-enforcing equilibria in coordination games and "institutions" enforcing "profitable" solutions in Prisoner's Dilemma and other mixed-motive games.

17. They will often be better represented by "game-learning" models than by "rational-analytic" models.

18. March and Olson (1989) opposed the "logic of appropriateness" to the "logic of consequentiality" rather than to a "logic of self-interest." But since not all norm-oriented behavior is purely expressive, one might better use two dimensions, "instrumental versus expressive" and "self-interested versus norm-oriented."

19. For Simmel (1984), this constituted the "sociological tragedy":

> Der Einzelne mag noch so feine, hochentwickelte, durchgebildete Qualitäten besitzen—gerade je mehr das der Fall ist, desto unwahrscheinlicher wird die Gleichheit und also die Einheitsbildung gerade dieser mit den Qualitäten anderer, desto mehr strecken sie sich nach der Dimension der Unvergleichbarkeit hin, auf desto niedrigere, primitiv sinnlichere Schichten reduziert sich das, womit er sich mit Sicherheit den anderen angleichen und mit ihnen eine einheitlich charakterisierte Masse formen kann. (P. 38)

20. The phenomenon has been discussed under a variety of labels, including "social motives" (MacCrimmon and Messick 1976), "styles of decision making" (Scharpf 1988, 1989), "social orientations" (Schulz and May 1989), or "social payoffs" (Burns 1989). The term proposed here has the advantage of emphasizing interaction effects and differing more clearly from "social norms."

21. For the complete list of experimentally identifiable "social orientations," see Schulz and May (1989).

REFERENCES

Alchian, Armen A. 1950. Uncertainty, evolution, and economic theory. *Journal of Political Economy* 58:211-21.

Aumann, Robert J. 1976. Agreeing to disagree. *Annals of Statistics* 4:1236-39.

———. 1987. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica* 55:1-18.

Axelrod, Robert. 1984. *The evolution of cooperation*. New York: Basic Books.

———. 1986. An evolutionary approach to norms. *American Political Science Review* 80:1095-1111.

Barber, Bernard. 1983. *The logic and limits of trust*. New Brunswick, NJ: Rutgers University Press.

Binmore, K. G. 1987. Why game theory "doesn't work." In *Analysing conflict and its resolution*, edited by P. G. Bennett, 23-42. Oxford: Clarendon.

Brams, Steven J. 1985. *Superpower games*. New Haven, CT: Yale University Press.

Brennan, Geoffrey, and James M. Buchanan. 1985. *The reason of rules: Constitutional political economy*. Cambridge: Cambridge University Press.

Buchanan, James M., and Gordon Tullock. 1962. *The calculus of consent: Logical foundations of constitutional democracy*. Ann Arbor: University of Michigan Press.

Burns, Tom R. 1989. *Notes: Toward a sociological theory of games and social behavior*. Manuscript, Uppsala University.

Burns, Tom R., and Helena Flam. 1987. *The shaping of social organizations: Social rule system theory with applications*. London: Sage.

Coleman, James S. 1964. *Introduction to mathematical sociology*. New York: Free Press.

Colman, Andrew M. 1982. *Game theory and experimental games*. Oxford: Pergamon.

David, Paul A. 1985. Clio and the economics of QUERTY. *American Economic History* 75:332-37.

Denzau, Arthur, William Riker, and Kenneth Shepsle. 1985. Farquharson and Fenno: Sophisticated voting and the home style. *American Political Science Review* 79:1117-34.

Dupy, Jean-Pierre. 1989. Common knowledge, common sense. *Theory and Decision* 27:37-62.

Eichhorn, Wolfgang. 1982. Contribution to Reinhard Selten. Einführung in die Theorie der Spiele mit unvollständiger Information. In *Information in der Wirtschaft*, edited by Erich Streißler, 149-50. Berlin: Duncker & Humblot.

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

————. 1983. *Explaining technical change: A case study in the philosophy of science*. Cambridge: Cambridge University Press.

————. 1989. *Nuts and bolts for the social sciences*. Cambridge: Cambridge University Press.

Etzioni, Amitai. 1988. *The moral dimension: Toward a new economics*. New York: Free Press.

Frank, Robert H. 1988. *Passions within reason: The strategic role of the emotions*. New York: Norton.

Friedman, James W. 1986. *Game theory with applications to economics*. Oxford: Oxford University Press.

Friedman, Milton. 1953. *Essays in positive economics*. Chicago: University of Chicago Press.

Fudenberg, Drew, and Eric Maskin. 1986. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica* 54:533-54.

Granovetter, Mark. 1985. Economic action and social structure: The problem of embeddedness. *American Journal of Sociology* 91:481-510.

Güth, Werner, Wolfgang Leininger, and Gunter Stephan. 1988. On supergames and folk theorems: A conceptual discussion. *Game Theory in the Behavioral Sciences* no. 19. Universität Bielefeld: Zentrum für Interdisziplinäre Forschung.

Guyer, Melvin, and Anatol Rapoport. 1969. Information effects in two mixed-motive games. *Behavioral Science* 14:467-82.

Hamburger, Henry. 1979. *Games as models of social phenomena*. San Francisco: Freeman.

Hannan, Michael T., and John Freeman. 1977. The population ecology of organizations. *American Journal of Sociology* 82:929-64.

————. 1984. Structural inertia and organizational change. *American Sociological Review* 49:149-64.

Hardin, Russell. 1980. The emergence of norms. *Ethics* 90:575-87.

————. 1988. *Morality within the limits of reason*. Chicago: University of Chicago Press.

Harsanyi, John C. 1967-68. Games with incomplete information played by Bayesian players. *Management Science* 14:159-82, 320-34, 486-502.

————. 1977. *Rational behavior and bargaining equilibrium in games and social situations.* Cambridge: Cambridge University Press.

Harsanyi, John C., and Reinhard Selten. 1988. *A general theory of equilibrium selection in games.* Cambridge: MIT Press.

Hayek, Friedrich A. 1960. *The constitution of liberty.* Chicago: University of Chicago Press.

Heiner, Ronald A. 1983. The origin of predictable behavior. *American Economic Review* 73:560-95.

Hernes, Gudmund. 1976. Structural change in social processes. *American Journal of Sociology* 82:513-47.

Jervis, Robert. 1978. Cooperation under the security dilemma. *World Politics* 30:167-214.

Kahn, Robert L., and Roderick M. Kramer. 1990. Untying the knot: Deescalatory processes in international conflicts. In *Nations and organizations: Contributions of organizational research to international relations,* edited by Robert L. Kahn and Mayer N. Zald. San Francisco: Jossey-Bass.

Kahneman, Daniel, and Amos Tversky. 1984. Choices, values, and frames. *American Psychologist* 39:341-50.

Kelley, Harold H. 1983. The situational origins of human tendencies. *Personality and Social Psychology Bulletin* 9:8-30.

Kelley, Harold H., and John W. Thibaut. 1978. *Interpersonal relations: A theory of interdependence.* New York: Wiley.

Kliemt, Hartmut. 1986. *Antagonistische Kooperation: Elementare spieltheoretische Modelle spontaner Ordnungsentstehung.* Freiburg: Alber.

Kliemt, H., and B. Schauenberg. 1984. Coalitions and hierarchies: Some observations on the fundamentals of human cooperation. In *Coalitions and collective action,* edited by Manfred J. Holler, 9-32. Wuerzburg: Physica.

Kreps, David M., and Robert Wilson. 1982. Reputation and imperfect information. *Journal of Economic Theory* 27:253-79.

Latsis, Spiros. 1972. Situational determinism in economics. *British Journal for the Philosophy of Science* 23:207-45.

Lindenberg, Siegwart. 1989. Choice and culture: The behavioral basis of cultural impact on transactions. In *Social structure and culture,* edited by Hans Haferkamp, 175-200. Berlin: de Gruyter.

Luhmann, Niklas. 1986. *Ökologische Kommunikation.* Opladen: Westdeutscher Verlag.

MacCrimmon, Kenneth R., and David M. Messick. 1976. A framework for social motives. *Behavioral Science* 21:86-100.

Macy, Michael W. 1989. Walking out of social traps. *Rationality and Society* 1:197-219.

March, James G., and Johan Olsen. 1989. *Rediscovering institutions: The organizational basis of politics.* New York: Free Press.

Maynard Smith, John. 1982. *Evolution and the theory of games.* Cambridge: Cambridge University Press.

Mayntz, Renate, and Thomas P. Hughes, eds. 1988. *The development of large technical systems.* Frankfurt: Campus.

McGinnis, Michael D. 1986. Issue linkage and the evolution of international cooperation. *Journal of Conflict Resolution* 30:141-70.

Milgrom, Paul R., Douglass C. North, and Barry R. Weingast. 1988. *Third-party enforcement of norms and contracts: A theoretical-historical analysis.* Paper prepared for the workshop on Connected Games, Max Planck Institut für Gesellschaftsforschung, Cologne, 17-18 November.

Molm, Linda D. 1989. Punishment power: A balancing process in power-dependence relations. *American Journal of Sociology* 94:1392-1418.

Neisser, Hans. 1952. The strategy of expecting the worst. *Social Research* 19:346-63.

Nelson, Richard R., and Sidney G. Winter. 1982. *An evolutionary theory of economic change.* Cambridge, MA: Belknap.

Ordeshook, Peter C. 1986. *Game theory and political theory: An introduction.* Cambridge: Cambridge University Press.

Ostrom, Elinor. 1986. An agenda for the study of institutions. *Public Choice* 48:3-25.

————. 1989. Microconstitutional change in multiconstitutional political systems. *Rationality and Society* 1:11-50.

Ostrom, Vincent. 1987. *The political theory of a compound republic: Designing the American experiment.* 2d ed. Lincoln: University of Nebraska Press.

Pagden, Anthony. 1988. The destruction of trust and its economic consequences in the case of eighteenth-century Naples. In *Trust: Making and breaking cooperative relations,* edited by Diego Gambetta, 127-41. Oxford: Blackwell.

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

Rasmusen, Eric. 1989. *Games and information: An introduction to game theory.* Oxford: Blackwell.

Ryll, Andreas. 1989. *Die Spieltheorie als Instrument der Gesellschaftsforschung.* MPIFG Discussion Paper 89/10. Köln: Max Planck Institut für Gesellschaftsforschung.

Scharpf, Fritz W. 1987. A game-theoretical explanation of inflation and unemployment in Western Europe. *Journal of Public Policy* 7:227-58.

————. 1988. The joint decision trap: Lessons from German federalism and European integration. *Public Administration* 66:239-78.

————. 1989. Decision rules, decision styles, and policy choices. *Journal of Theoretical Politics* 1:149-76.

Schotter, Andrew. 1981. *The economic theory of social institutions.* Cambridge: Cambridge University Press.

Schulz, Ulrich, and Theo May. 1989. The recording of social orientations with ranking and pair comparison procedures. *European Journal of Social Psychology* 19:41-59.

Selten, Reinhard. 1978. The chain-store paradox. *Theory and Decision* 9:127-59.

————. 1982. Einführung in die Theorie der Spiele mit unvollständiger Information. In *Information in der Wirtschaft,* edited by Erich Streißler, 81-147. Berlin: Duncker & Humblot.

————. 1985. Comment. In *Frontiers of economics,* edited by Kenneth J. Arrow and Seppo Honkapohja, 77-87. Oxford: Blackwell.

Shepsle, Kenneth A. 1979. Institutional arrangements and equilibrium in multidimensional voting models. *American Journal of Political Science* 23:27-59.

————. 1989. Studying institutions: Some lessons from the rational choice approach. *Journal of Theoretical Politics* 1:131-47.

Shepsle, Kenneth A., and Barry R. Weingast. 1981. Political preferences for the pork barrel: A generalization. *American Journal of Political Science* 25:96-111.

————. 1987. The institutional foundations of committee power. *American Political Science Review* 81:85-104.

Simmel, Georg. [1917] 1984. *Grundfragen der Soziologie, Individuum und Gesellschaft.* 4th ed. Berlin: de Gruyter.

Snyder, Glenn H. 1971. "Prisoners' Dilemma" and "Chicken" models in international politics. *International Studies Quarterly* 15:66-103.

Sugden, Robert. 1989. Spontaneous order. *Journal of Economic Perspectives* 3:85-97.

Taylor, Michael. 1987. *The possibility of cooperation*. Cambridge: Cambridge University Press.

Thompson, Earl A., and Roger L. Faith. 1980. Social interaction under truly perfect information. *Journal of Mathematical Sociology* 7:181-97.

————. 1981. A pure theory of strategic behavior and social institutions. *American Economic Review* 71:366-80.

Tullock, Gordon. 1985. Adam Smith and the Prisoners' Dilemma. *Quarterly Journal of Economics*, Supplement:1073-81.

Ullmann-Margalit, Edna. 1977. *The evolution of norms*. Oxford: Clarendon.

Witt, Ulrich. 1986. Evolution and stability of cooperation without enforceable contracts. *Kyklos* 39:245-66.

————. 1989. The evolution of economic institutions as a propagation process. *Public Choice* 62:155-72.

Zintl, Reinhard. 1989. Der Homo Oeconomicus: Ausnahmeerscheinung in jeder Situation oder Jedermann in Ausnahmesituationen? *Analyse & Kritik* 11:52-69.