Elinor Ostrom

Collective Action and Local Development Processes

(doi: 10.2383/25950)

Sociologica (ISSN 1971-8853)
Fascicolo 3, novembre-dicembre 2007

Ente di afferenza:
()

Collective Action and Local Development Processes

by Elinor Ostrom

doi: 10.2383/25950

The theory of collective action focuses on the question of why individuals would cooperate in a social dilemma when they could free-ride on the contributions of others. The term "social dilemma" refers to a setting in which individuals choose actions in an interdependent situation. If each individual in such situations selects strategies based on a calculus that maximizes short-term benefits to self, individuals are predicted to take actions that generate lower joint outcomes than could have been achieved. In other words, a social dilemma can be analyzed as a game where the predicted equilibrium for a single iteration of the game yields less than the socially optimal outcome. The socially optimal outcome could be achieved if those involved "cooperated" by selecting strategies other than those prescribed by game theory. Since the suboptimal joint outcome is an equilibrium, no one is independently motivated to change their choice, given the predicted choices of all others.

Such situations are *dilemmas* because at least one outcome yields higher returns for *all* participants, but rational participants making independent choices are predicted not to achieve this outcome. Social dilemmas involve a conflict between individual rationality and optimal outcomes for a group [Schelling 1978; Lichbach 1995]. Even if some individuals cooperate, others are predicted to "free-ride" on the contributions of the cooperators.

In addition to the assumption regarding the structure of payoffs leading to a deficient equilibrium, further assumptions made in almost all models of social dilemmas include:

1. Decisions about strategies are made independently and simultaneously;

2. all participants have common knowledge of the exogenously fixed structure of the situation and of the payoffs to be received by all individuals under all combinations of strategies;

3. no external actor (or central authority) is present to enforce agreements among participants about their choices.

When these assumptions are made for a game that is repeated only once, the theoretical prediction derived from noncooperative game theory is unambiguous – zero cooperation.

In a very large number of one-shot public good experiments undertaken in diverse countries, however, subjects tend to contribute an average amount between 40 to 60 percent of the optimal level of contributions [Davis and Holt 1993: 325; Sally 1995].[1]

If such a game is finitely repeated, and everyone shares complete information about the structure of the situation, the predicted outcome of noncooperative game theory for each iteration of the game is again the equilibrium of the constituent game. The presumption is that if individuals would not cooperate in the last game of a series, they would not cooperate in the second-to-last game, the third-to-last game, all the way back to the initial game [Luce and Raiffa 1957]. Again, extensive experimental research has shown that cooperation in finitely repeated social dilemmas is much higher than predicted using backward induction while it does decay over time as the last iteration of the series is approached [see E. Ostrom 1998; 2005, for discussion of these extensive research findings]. Amnon Rapoport [1997: 122] has gone so far as to assert that "subjects are not involved in or capable of backward induction."

When uncertainty exists about the time or the number of rounds involved in a repeated game (or, if the repetition is infinite), two theoretical developments generate more optimistic predictions than backward induction in finitely repeated games. First, Kreps *et al.* [1982] posited that if some individuals in a game do *not* follow the prescriptions of full rationality involving the maximization of expected objective outcomes to self, other fully rational players might then adopt cooperative strategies at least in the early stages of a game so as to gain the benefits of engaging in reciprocal cooperation. Second, Fudenberg and Maskin [1986] posited that it was possible for subjects to eliminate free riding if some players made a firm commitment to follow

[1] The concern of some scholars about the validity of experiments that have been conducted with college students that are paid modest sums has been addressed by several studies that have greatly increased the monetary value of the payoffs offered to subjects [Cameron 1999]. Studies that recruited workers [Carpenter, Burks and Verhoogen 2005] or rural villagers [Cardenas 2000] instead of students as subjects also have not shown significant differences in the patterns of responses [see also Henrich *et al.* 2004].

a "grim trigger strategy." A grim trigger strategy involves a permanent switch from cooperation to defection once anyone fails to cooperate. This self-enforcing, positive equilibrium is only possible if all players strongly commit themselves to punish others and deter defection by their known strong commitment.

These theoretical results have held up over the years. Instead of generating a clear and better prediction, however, they lead to an explosion of the number of possible equilibria predicted by noncooperative game theory. Among the predicted equilibria are strategies yielding the suboptimal equilibrium, the optimal outcome, and everything in between [Abreau 1988]. Thus, while empirical evidence generates some optimism that collective action can be achieved in some settings, the *problem* of collective action remains: How can participants in social dilemmas avoid the temptation of suboptimal equilibria and move closer to optimal outcomes – in other words, gain a "cooperators' dividend" [Lichbach 1996].

Developing a coherent theory of collective action that is relevant for practice in explaining development is a real challenge. At the individual level, individuals do take costly actions that effectively take the interests of others into account. Shivakumar [2005] and Gellar [2005] provide evidence of local and regional groups that are successfully engaging in collective action in Somaliland and in Senegal where little cooperation occurred earlier. On the other hand, individuals may callously ignore or viciously harm others depending on the setting in which they find themselves [see Fiske, Harris and Cuddy 2004]. "For example, in many societies, predatory activities that would be cause for shame when directed at one's own group, are cause for pride when strangers are victimized. The formidable task confronting any comprehensive account of social cooperation is to explain both the remarkable robustness of cooperation, and the pattern underlying its lapses" [Heckathorn 1991: 2].

Thus, an important task for all social scientists is achieving a more coherent synthesis of theoretical work that posits variables affecting the likelihood of undertaking diverse forms of collective action. We must be able to explain success as well as failure of efforts to achieve collective action. Further, we need to recognize that forms of collective action differ in regard to the distribution of benefits and harms to those in a group and those who are external to it. Mobs, gangs, and cartels are forms of collective action as well as neighborhood associations, charities, and voting.

In this paper I will first discuss the growing and extensive theoretical literature positing a host of structural variables presumed to affect the likelihood of individuals achieving collective action to overcome social dilemmas. None of these structural variables, however, should really make any difference in the probability of successful collective action if we continue to treat the *model* of rationality that has proved successful in explaining behavior and outcomes in competitive market settings as a universal *theory* of human behavior. Thus, in the second section I will examine how a theory of boundedly rational, norm-based human behavior is a better foundation for explaining collective action than a model of maximizing material payoffs to self. If one posits that individuals can use reciprocity and reputations to build trust in dilemma situations, then one can begin to explain *both* successful and unsuccessful efforts to overcome social dilemmas through collective action.

The third section will then briefly examine the linkage between the structural measures discussed in the first section with the core individual relationships discussed in the second. The fourth section will introduce the topic of changing the rules of a focal dilemma in deeper arenas in efforts to improve the net benefits from collective action by affecting the structural variables of the focal arena. In conclusion, I will reflect on the challenge that social scientists face in testing collective-action theory in light of the large number of variables posited to affect outcomes.

Structural Variables Predicted to Affect the Likelihood of Collective Action

A rich array of theoretical speculations, formal game-theoretic models, and computer models of evolutionary processes have generated a long list of structural variables that are frequently postulated to affect the likelihood that a set of participants will be able to achieve outcomes greater than the deficient equilibrium – or, the cooperators' dividend [Lichbach 1996]. Let us first focus on structural variables that do not essentially depend on a situation being repeated. These include:

1. the number of participants involved;
2. whether benefits are subtractive or fully shared (i.e., public goods vs common-pool resources);
3. the heterogeneity of participants;
4. face-to-face communication;
5. the shape of the production function.

Then, we will focus on situations where repetition of the situation makes possible the impact of additional structural variables including:

6. information about past actions;
7. how individuals are linked;
8. whether individuals can enter or exit voluntarily.

Let us turn to a brief discussion of these eight major variables.

Situations Where Repetition is Not Relevant

Among the variables posited to affect the likelihood of participants overcoming a one-shot social dilemma are the number of participants, whether benefits are subtractive or fully shared, their heterogeneity, whether they can communicate, and the shape of the production function they face [E. Ostrom 2001].

The Number of Participants Involved

In his influential book *The Logic of Collective Action*, Mancur Olson [1965] argued that as the size of a group increased, the probability of a group achieving a public good decreased and the extent of nonoptimality increased. Olson posited two reasons for this hypothesis. First, as group size increases, the noticeability of any single input to the provision of a public good declines. It is then easier for individuals to think that their own free riding will not be noticed and thus not affect the likelihood that the good will be provided. Second, coming to an internal agreement about coordinated strategies in larger groups involves higher transaction costs. Thus, a core theoretical hypothesis has been that the number of participants will likely reduce the probability of achieving any form of collective action or at least diminish the amount of joint benefits that could be achieved.

On the other hand, some theorists have generated the opposite prediction from those based on the work of Olson [1965]. In an effort to understand the phenomenon of age grade organization which were so frequently used in most of Africa as a means of providing public goods – particularly defense – Bates and Shepsle [1995] developed a formal model of a three-period, overlapping generations, public good game. A corollary of this model generates a prediction that the provision of public goods is *positively* correlated with group size. Agrawal [2000] posits a curvilinear relationship between size of group and collective action.

The impact of group size has been subject to considerable theoretical debate. Chamberlin [1974] pointed out that differences in group size frequently affect other key variables including the marginal impact of an individual's contribution of a fixed amount [see also Frohlich and Oppenheimer 1970; Pecorino 1999]. Thus, how size might affect the likelihood of cooperation depends on how other structural variables are affected by the size of a group.

Subtractive versus Fully-Shared Benefits

In Olson's original analysis, he included all dilemmas where it was difficult to exclude potential beneficiaries, whether or not they had contributed. Unfortunately, Olson's analysis confounded situations where the consumption of benefits by one individual subtracted benefits from others with situations where consumption was nonsubtractive in nature [characterized as having full jointness of supply: see Ostrom and Ostrom 1999]. In a public good environment, increasing the number of participants tends to bring additional resources that could be drawn on to provide a benefit that will be jointly enjoyed by all. It is because of the additional resources available in a larger group and the nonsubtractability characteristic of public goods, that Marwell and Oliver [1993: 45] conclude that when "a good has pure jointness of supply, group size has a *positive* effect on the probability that it will be provided."

Goods that are subtractable in nature are better defined as common-pool resources (CPRs) [E. Ostrom, Gardner, and Walker 1994]. Social dilemmas related to CPRs share with public good provision the problems of free riding, but they also include the problems of overharvesting and crowding. In a CPR environment, an increase in the number of participants, holding other variables constant, is negatively related to achieving social benefits.

Weissing and Ostrom [1991] analyzed a formal game examining the impact of the number of individuals involved in a CPR game where each player has an opportunity to take a legal amount of water from an irrigation system or steal water and between monitoring or not the behavior of others in the system. When all other variables are analytically held constant, an increase in the number of players increases the rate of stealing at equilibrium. However, many variables are affected by increasing the number of participants. The value of water at the margin for irrigators is likely to increase (thus making stealing more attractive). The impact of one person's stealing may be spread out over more individuals and thus the loss to any one farmer of someone else stealing water may be less severe at the margin (thus making monitoring less attractive). An increase in the number of participants may also mean a larger system where more water is available and the consequences listed above would then not follow. Thus, in a CPR environment, whether size has a positive impact, a negative impact, or any impact is dependent upon how other variables are affected by a change in the number of participants.

The Heterogeneity of Participants

Participants can be heterogeneous in many ways. Olson [1965] argued that if there were one or a few individuals who had much stronger interests in achieving a public good (in other words, they faced different payoff functions), the probability of a group achieving a public good increased even though the good was still likely to be underprovided.[2] Others have speculated that heterogeneity in assets, information, and payoffs are negatively related to gaining a cooperators' dividend due principally to increased transaction costs and the conflict that would exist over the distribution of benefits and costs to be borne. In fact, the literature contains many arguments that point to heterogeneity as a serious deterrent to cooperation [Hardin 1982; Johnson and Libecap 1982; Libecap and Wiggins 1984; Isaac and Walker 1988; Wiggins and Libecap 1987; Bardhan 1993; Seabright 1993]. E. Jones [2004] reasons that the presence of wealthy participants may encourage trust in them early in a process of collective action and encourage the formation of cooperatives. Inequality in distribution of benefits may, however, reduce trust and cooperation later in the process. The impact of heterogeneity on levels of collective action achieved frequently interacts with the shape of the production function for a good and thus will be discussed further below.

Face-to-Face Communication

Given that noncooperative game theory predicts that communication will make *no* difference in the outcome of social dilemmas, the repeated findings of a strong positive effect that communication has on the outcomes of collective-action experiments is a major theoretical puzzle [Sally 1995]. The result has been replicated so many times, however, that contemporary scholars have to take it seriously.

Frohlich and Oppenheimer [1998] explain the effectiveness of communication in general related to the needs of individuals in such settings to express the desire to each other that they should forego their immediate self-interest for the benefit of the group. In other words, communication is used for "moral suasion." And, being able to look others directly in the eye while discussing such moral issues is substantially better than relying on written communication. Kerr and Kauf-

[2] Closely related to the concept of a privileged group is the international relations theory of hegemonic stability [Keohane 1984] that posits that heterogeneity promotes cooperation because large actors are endowed with more resources (including the power to coerce others) and are better able to produce a public good such as international peace. The theory predicts that when there are a limited number of larger states dominating international relations that the collective good of peace is more likely to be provided.

man-Gilliland [1994] conclude that communication in general helps a group gain a sense of "solidarity" and that face-to-face communication enhances the likelihood that individuals will keep their promises to cooperate. In general, the efficacy of communication appears to be related to the increased trust that individuals acquire when promises are made to them in a face-to-face setting. When they are in a repeated situation, they use the opportunity for communication to discuss deviations from promises made in a highly critical and moralistic tone [E. Ostrom, Gardner and Walker 1994; Parks, Henager and Scamahorn 1996; Valley, Moag and Bazerman 1998].

The Shape of the Production Function

All social dilemmas involve individuals who could take actions that produce benefits for others (and themselves) at a cost that they themselves must bear. The production function that relates individual actions to group outcomes may take any of a wide diversity of forms. One possible form is a step function (*b* in Figure 1), in which actions by up to *k* participants make no difference in the benefit function, but actions by *k* or more participants discontinuously shift the benefit functions upward. Hardin [1976] was among the first to argue that when the shape of the production function for a public good was a step function, solving social dilemmas would be facilitated since no good would be provided if participants did not gain sufficient inputs to equal or exceed the provision point (*k*). Until the benefit is actually produced, it is not possible to "free ride" on the contribution of others. In these settings, individuals may assume that their participation is critical to the provision of the good. This type of production function may create an "assurance problem" rather than a strict social dilemma. For those who perceive their contribution as critical, not contributing is no longer the unique equilibrium. Participants are motivated to contribute so long as they expect that other critical participants will contribute.

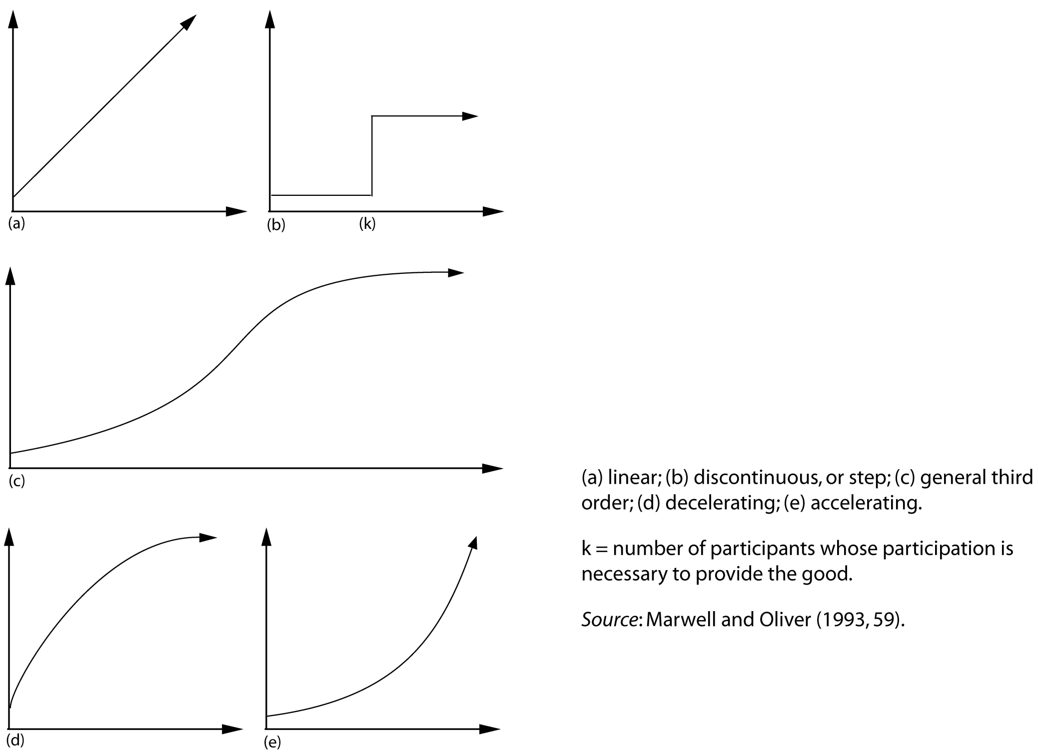

FIG. 1. General types of production function.

Closely related to this attribute of the production function itself, are sharing formulas that may be developed by participants to make each person of the entire group, or a designated minimal contributing group, feel that their contribution is critical [van de Kragt, Orbell and Dawes 1983]. By agreeing that each person will contribute a set proportion of what is believed to be the total cost of obtaining a good, the individuals in such a minimal contributing set face a choice between contributing and receiving the benefit (assuming others in the minimal contributing set also contribute), or not contributing and receiving nothing.

Strict step functions or discrete goods are relatively unusual production functions. Marwell and Oliver [1993] conduct an extensive analysis of monotonically increasing, linear and nonlinear production functions relating individual contributions and the total benefits produced. Linear production functions are used extensively in N-person PD and public good games where the prediction is that a homogeneous group will contribute zero resources. Marwell and Oliver focus on nonlinear functions and distinguish between third-order production functions that are decelerating and those that are accelerating. In the decelerating case (*d* in Figure 1), while every contribution increases the total benefits that a group receives, marginal returns decrease as more and more individuals contribute. When contributions are made sequentially, the initial contributions have far more impact than later contributions.

The example they use to illustrate such a production function is calling about a pothole in a neighborhood where a city administration is sensitive to citizen support [*ibidem*: 62]. The first call brings the pothole to the attention of city officials and puts it on the list of things to be repaired (raising the probability of repair from zero to perhaps .4 or higher). The second call increases the probability of repair still further, but not as much as the first call. Later calls continue to increase the probability but with a smaller and smaller increment.

With an accelerating production function (*e* on Figure 1), initial contributions make small increments and later contributions yield progressively greater benefits. "Accelerating production functions are characterized by *positive interdependence*: each contribution makes the next one more worthwhile and, thus, more likely" [*ibidem*: 63]. Settings where mass actions are needed in order to gain a positive response involve accelerating functions. A strike involving only a few workers is unlikely to produce the level of benefits yielded by a strike involving a very large proportion of the workers of a firm or in an industry.

The theoretical predictions depend sensitively on the particular shape of the production function, on whether all participants are symmetric or have different levels of assets, on the sequence in which individuals contribute, and on the information generated by each action. For homogeneous groups facing decelerating curves, which Marwell and Oliver assert characterize many field situations involving large numbers of potential beneficiaries, getting over the initial period where the returns to participants are negative defeats collective action before it can generate sufficient inputs to gain net benefits. Thus, collective goods that have a decelerating production function are unlikely to be provided by large groups of relatively homogeneous individuals acting independently, or if provided, they will be provided as Olson predicted at a suboptimal level. The prediction for homogeneous groups and accelerative functions is also gloomy. The key is whether the initial contributions are made and this is somewhat less likely with a homogeneous group than with a heterogeneous group who may have some members with high levels of interest and who would be more interested in contributing the initial inputs.

Heterogeneous groups facing accelerative production functions may need substantial organization to overcome the initial start-up costs. Whether these benefits are produced at all depends on the presence of a *critical mass* of individuals (a subgroup of individuals who have a sufficient large interest in the benefits received and sufficient resources needed to cover the initial start-up costs). In other words, some degree of heterogeneity of the valuation of joint outcomes, and in many cases of the resources needed to generate those outcomes, is necessary to achieve the cooperators' dividend. If interest in joint outcomes and availability of resources are positively correlated, the likelihood of there being a critical mass is higher [*ibidem*: 87]. Once the initial contributions are made, a bandwagon effect may occur where those with a lower valuation of the outcome can contribute and see a substantial return on their contribution. Marwell and Oliver repeatedly discuss the importance of participants communicating with one another to cope with the diverse problems associated with heterogeneity of interest combined with the shape of the production function, but communication is not included in their formal analysis.

Repetition of Interactions

With repeated interactions, at least three more structural variables are posited to affect the level of cooperation achieved in social dilemma situations: the level of information generated about past actions, how individuals are linked, and voluntary entry and exit.

Information about Past Actions

In a two-person game where individuals know the structure of the game and learn accurate information about the outcomes achieved, the behavior of the other individual is also known. As soon as more than two individuals are involved, accurate information about outcomes alone is no longer sufficient to inform one player about the actions of others. In families and small neighborhoods, where interactions are repeated, reputations can be built over time and group members can build up a level of trust about other participants [Seabright 1993]. Cooperation can grow over time in such settings. In large groups, the disjunction between an individual's actions and reputations is more difficult to overcome. In some situations, individuals can observe the actions of others and thus know what each individual did in the previous rounds. Various ways of monitoring the actions of participants increase or decrease the availability and accuracy of the information that individuals have concerning the particular actions of known individuals (or types of players) in the past [Janssen 2004].

How Individuals are Linked

Sociologists and social psychologists have stressed the importance of how individuals may or may not be linked in a network when confronting various types of social dilemmas [Granovetter 1973; Cook and Hardin 2001]. They have posited that individuals who are linked in a network where A contributes resources to B, and B contributes resources to C, and C contributes resources to A – or any similar unidirectional linking – are more likely to contribute to each other's welfare than individuals whose resource contribution goes to a generalized pool from which all individuals obtain benefits. The reason given for this expectation is that individuals in an undifferentiated group setting can expect to free ride for a longer period of time without reducing their own benefits than when contributions have to be delivered to someone in the chain of relationships in order for benefits to eventually come to them. Anyone in the chain who stops contributing faces a higher probability (so the argument goes) of the chain of benefit-enhancing contributions stopping and their losing out on obtaining a positive benefit. Creating a particular type of network may change the structure of the game from an n-person PD to an Assurance Game [Yamagishi and Cook 1993].

The Possibility of Choosing Whether to Play or Not (Entry and Exit)

Orbell and Dawes [1991] and Hauk and Nagel [2001] have argued that when individuals have a choice as to whether to play social dilemma games with others, and they can identify the individuals with whom they have played, that individuals will choose partners so as to increase the frequency with which cooperative outcomes are achieved. This gives individuals a third choice in a social dilemma game. Besides deciding whether to cooperate, they can decide whether to "opt out." If one player opts out, the decision round ends, and everyone receives a zero payoff. All players have an effective veto over the entire play of the game.

Janssen [2005] has developed an agent-based model of a two-person, prisoner's dilemma in which individuals can cooperate, defect or withdraw. Each agent carries symbols that can be identified by others. The symbols are used by participants to gain or lose trust that the other participant will cooperate. Given this capacity to recognize trustworthiness in others *and* the capacity to withdraw from playing a game at all, cooperation levels rise over time and reach relatively high levels in populations composed of 100 players. With 1,000 players, cooperation levels are lower unless the number of symbols that can be used to recognize trustworthy plays is increased – a somewhat counterintuitive result [see also Hauert *et al.* 2002].

Towards a More General Theory of Human Behavior

The assumption that individuals are inexorably stuck within social dilemmas has slowly been replaced in some theoretical work with a recognition that individuals face the *possibility* of achieving results that avoid the worst outcomes and, in some situations, may even approximate optimality. The clear and unambiguous predictions of earlier theories have been replaced with a broad range of predictions including some that are far more optimistic. The theoretical enterprise has, however, become more opaque and confused.

This is a particularly challenging puzzle for scholars who yearn for frameworks and theories of behavior that integrate across the social sciences. To have one theory – rational choice theory – that explains how individuals achieve close to optimal outcomes in markets, but fails to explain why anyone votes or contributes voluntarily to the provision of public goods, is not a satisfactory state of knowledge in the social sciences. Simply assuming that individuals are successfully socialized into seeking better group outcomes does not explain the obvious fact that groups often fail to obtain jointly beneficial outcomes [Dietz, Ostrom, and Stern 2003].

We need to recognize that what has come to be called rational choice *theory* is instead one *model* in a family of models useful for conducting formal analyses of human decisions in highly structured settings. It is a thin model of a broader theory of rational behavior. When it is used successfully, the rational choice model is largely dependent for its power of explanation on how the structure of the situations involved is modeled [Satz and Ferejohn 1994]. In other words, the context within which individuals face social dilemmas is more important in explaining levels of collective action than relying on a single model of rational behavior as used in classical noncooperative game theory [see Orbell *et al.* 2004].

In highly structured and competitive environments, predictions generated from the *combination* of a model of the situation and a model of complete rationality are well-supported empirically. As Alchian [1950] demonstrated long ago, competitive markets eliminate businesses that do not maximize profits. Further, markets generate limited, but sufficient, statistics needed to maximize profits. The institutional structure of a market rewards individuals who make economically rational decisions and who can then be modeled as if they were determinate, calculating machines.

A broader theory of human behavior views humans as adaptive creatures [B. Jones 2001] who attempt to do as well as they can given the constraints of the situations in which they find themselves [Simon 1955; 1999]. Humans learn norms, heuristics, and full analytical strategies from one another, from feedback from the world, and from their own capacity to engage in self-reflection and imagine a differently structured world. They are capable of designing new tools – including institutions – that can change the structure of the worlds they face for good or evil purposes. They adopt both short-term and long-term perspectives dependent on the opportunities they face. Multiple models are consistent with a theory of boundedly rational human behavior, including a model of complete rationality when paired with repetitive, highly competitive situations.

Heuristics and Norms

In many everyday situations individuals tend to use heuristics – rules of thumb – that they have learned over time regarding responses that tend to give them good (but, not necessarily optimal) outcomes in particular kinds of situations. In frequently encountered, repetitive situations, individuals learn better and better heuristics that are tailored to the particular situation. With repetition and sufficiently large stakes, individuals may learn heuristics that approach best-response strategies and thus approach local optima [Gigerenzer and Selten 2001].

Many theorists interested in collective action have focused on the potentially positive effects of participants adopting simple heuristics to use when they are in a social dilemma situation. Morikawa, Orbell and Runde [1995], for example, examine the efficacy of using the simple heuristic of "expect others to have the same dispositions as yourself." They conduct a computer simulation where each actor in a population of 10,000 actors is matched to another actor. Those simulated actors whose payoff is above the mean are multiplied by two, while those whose payoff is below the mean are eliminated from the simulation. From their simulations, they predict that the heuristic is of most value to individuals who are moderately disposed to cooperate rather than holding either of the extremes. Their simulation also generates the prediction that the heuristic will be most valuable when social dilemmas occur among those in close proximity and that the probability of there being some very cooperative groups of agents increases with the size of the population.

In addition to learning instrumental heuristics, individuals also learn norms. By norms, I mean that the individual attaches an internal valuation – positive or negative – to taking particular types of action. Analytically, individuals can be thought of as learning norms of behavior that are relatively general and fit a wide diversity of particular situations. Crawford and Ostrom [2005] refer to this internal valuation as a delta parameter that is added to or subtracted from the objective costs of an action or an outcome. Andreoni [1989] models individuals who gain a "warm glow" when they contribute resources that help others more than they help themselves in the short term. Knack [1992] refers to negative internal valuations as "duty." The strength of the commitment [Sen 1977] made by an individual to take particular types of future actions (telling the truth, keeping promises), is reflected in the size of the delta parameter. After experiencing repeated benefits from their own and from other people's cooperative actions, individuals may resolve that they should always initiate cooperation in the future.[3] Or, after many experiences of being the "sucker" in such experiences, an individual may resolve never to initiate unilateral cooperation and to punish noncooperators whenever feasible.

James Cox and colleagues posit that individual behavior in a particular setting is affected by an individual's initial emotional or normative state and then by direct experience with others in a specific setting [Cox 2004; Cox and Deck 2005]. The underlying norms and direct experience in a particular setting combine to affect orientations toward reciprocity [Cox, Friedman and Gjerstad forthcoming].

Fairness is also one of the norms used by individuals in social dilemma settings. The maximal net return to a group may be obtained in a manner that is perceived to be fair or unfair by those involved. When participants are symmetric in regard to all strategically relevant variables, the only real fairness issue relates to the potential capability of some to free ride on others [Dawes, Orbell and van de Kragt 1986]. When participants differ, however, finding an allocation formula perceived by most participants as fair is far more challenging. In both cases, however, theorists have argued that when participants think that a proposal for sharing costs and benefits is fair, they are far more willing to contribute [Isaac, Mathieu and Zajac 1991].

Since norms are learned, they vary substantially across individuals, and within individuals across the different types of situations they face, and across time within any particular situation. As Brennan and Pettit [2004] stress, however, norms that help to solve social dilemmas need to be shared so that individuals who act contrary to the norm fear the reduction in esteem likely to occur. Once some members of a population acquire norms of behavior, they affect the expectations of others. When interacting with individuals who are known to use retribution against those who are not trustworthy, one is better off by keeping one's commitments.

Contingent Strategies and Norms of Reciprocity

Many theorists posit that one can explain behavior in social dilemmas better if one assumes that boundedly rational individuals enter situations with an initial probability of using reciprocity based either as a calculated strategy that contingent action

[3] Whenever games are repeated, the discount rates used by individuals also affects the adoption of norms including that of reciprocity. In settings where individuals do not strongly discount outcomes that will occur in the distant future, they can realize the benefits of cooperation over a long series of plays – thus offsetting the initial material advantage of not cooperating. As the future is more strongly discounted, however, the calculation made by an individual focuses more on the immediate materials payoffs.

leads one to be better off *or* based on a normative belief that this is how one should behave [Bolton and Ockenfels 2000; Falk, Fehr and Fischbacher 2002; Panchanathan and Boyd 2004]. In either case, individuals learn to use reciprocity based on their own prior training and experience. The more benefits that they have received in the past from other reciprocators, the higher their own initial inclinations. The more they have faced retribution, the less likely they estimate that free riding is an attractive option.

By and far the most famous contingent strategy – tit-for-tat – has been the subject of considerable study from an evolutionary perspective. In these analyses, pairs of individuals are sampled from a population who then interact with one another repeatedly in a PD game. Each individual is modeled as if they had inherited a strategy including the fixed maxims of always cooperate, always defect, or the reciprocating strategy of tit-for-tat (cooperate first, and then do whatever the others did on the last round). Axelrod and Hamilton [1981] and Axelrod [1984] have shown that when individuals are grouped so that they are more likely to interact with one another than with the general population, and when the expected number of interactions is sufficiently large, reciprocating strategies such as tit-for-tat can successfully invade populations composed of individuals following an all-defect strategy. But the size of the population in which interactions are occurring must be relatively small for reciprocating strategies to survive potential errors of players [Bendor and Mookherjee 1987]. Boyd and Richerson [1988] have examined a model where more than two individuals are sampled from a large population to interact repeatedly in an n-person prisoner's dilemma. They conclude that increasing the size of the relevant population reduces the probability that selection will favor reciprocating strategies unless tight subgroups are formed that rarely interact across subgroup boundaries.

Reciprocating strategies continue to limit what individuals can do who face others who do not cooperate. The only way of "punishing" defection is to defect oneself, which may lock participants into the deficient equilibrium. Punishment in field settings usually involves some action other than defecting oneself on an agreement. Since punishing someone else usually involves a cost for oneself and produces a benefit for everyone, it is a second-order social dilemma.

Boyd and Richerson [1992] build a two-stage evolutionary model of a large population from which groups of size $n > 2$ are selected. The first stage is an n-person PD where an individual selects cooperate or defect. In the second stage, any individual can punish any other individual at a cost to the punisher and to the punished. The same group continues for the next round dependent on a probability function. Strategies are modeled as if they were inherited. They allow errors to occur in the execution of a cooperative strategy, but all other strategies are executed as intended.

After the rounds of interaction are completed, the more successful strategies are reproduced at a higher rate than the less successful strategies.

In the Boyd and Richerson [1992] model, an increase in group size requires an offsetting linear increase in the number of interactions to achieve similar levels of collective action [see also Richerson and Boyd 2005]. They also find that moralistic strategies, "which punish defectors, individuals who do not punish noncooperators, and individuals who do not punish nonpunishers can also overcome the problem of second-order cooperation" [*ibidem*: 184]. When moralistic strategies are common, defectors and cooperators who do not punish are selected against due to the punishment directed at them. "In this way, selection may favor punishment, even though the cooperation that results is not sufficient to compensate individual punishers for its costs" [*ibidem*]. These moralistic strategies can stabilize any behavior – a result that is similar to the famous "folk theorem" that any equilibrium can be stabilized by such punishing strategies as the grim trigger.

Several of the heuristics or strategies posited to help individuals gain larger cooperators' dividends depend upon the willingness of participants to use retribution to at least some degree. In tit-for-tat, for example, an individual must be willing to "punish" a player who defected on the last round by defecting on the current round. In repeated games where substantial joint benefits are to be gained from mutual cooperation, the threat of the grim trigger is posited to encourage everyone to cooperate. A small error on the part of one player or exogenous noise in the payoff function, however, makes this strategy a very dangerous one to use in large environments where the cooperators' dividend is substantial.

Güth and Kliemt [1995] show that retributive emotions can survive in evolutionary stable ways if it is possible for players to know in advance whether the person with whom they are playing is characterized by a "strong conscience" or a willingness to impose punishments if cooperation is not selected. Using an indirect evolutionary approach in which preferences become endogenous, they show that including another in one's utility function depends on the favorable response of the other to cooperative moves. Family members, in particular, are more likely to have other family members in their utility functions, but their argument differs from the kin-selection argument. Further, the evolution of preferences that include benefits to others is more likely to emerge in populations where individuals are not anonymous and can use symbols to identify their type [Ahn, Janssen and Ostrom 2004].

The Core Relationships: Reputation, Trust, and Reciprocity as They Affect Cooperation

In situations where individuals can acquire a reputation for using positive and negative reciprocity and being trustworthy, others can learn to trust those with such a reputation. At the core of an evolving theoretical explanation of successful or unsuccessful collective action are the links between the trust that one participant (P_i) has in the others ($P_j \ldots P_n$) involved in a collective-action situation, the investment others make in trustworthy reputations, and the probability of all participants using reciprocity norms (see Figure 2). When some individuals initiate cooperation in a repeated situation, others learn to trust them and are more willing to adopt reciprocity themselves leading to higher levels of cooperation. And, when more individuals use reciprocity, gaining a reputation for being trustworthy is a good investment as well as an intrinsic value. Thus, reputations for being trustworthy, levels of trust, and reciprocity are positively re-enforcing. This also means that a decrease in any one of these can generate a downward cascade leading to little or no cooperation.

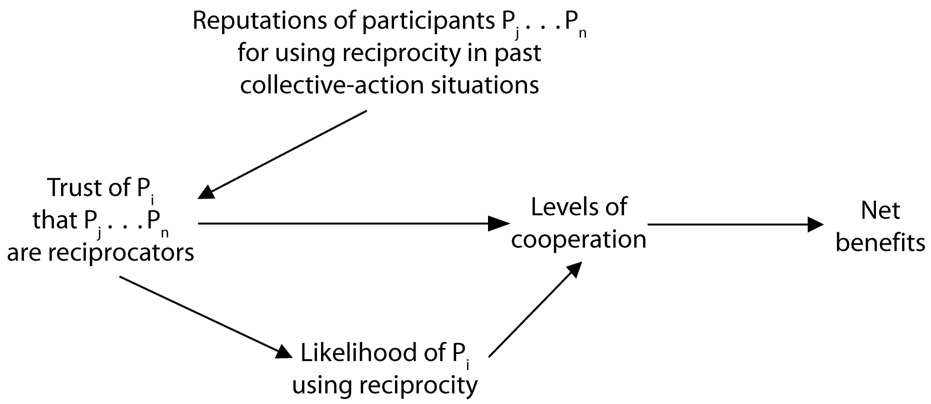

FIG. 2. The core relationships at the individual level affecting levels of cooperation in a social dilemma.

Linking Structural Variables to the Core Relationships

Instead of explaining cooperation directly from the material incentives facing individuals in social dilemmas, the task we face in explaining developmental processes is linking external structural variables to an inner core of individual level variables – reputation, trust, and reciprocity – as these in turn affect levels of cooperation and net benefits achieved. Some potential linkages are now pretty obvious. One can confidently posit that in a small, homogeneous group interacting in a face-to-face meeting to discuss producing a public good with an accelerating production function, the costs of coming to an agreement will tend to be low and the probability that individuals keep their promises will be high. Previous gossip will have identified which members of the group could be trusted to keep agreements and efforts to exclude such untrustworthy participants would be undertaken. The combined effect of the structural variables in this example on reputation, trust, and reciprocity is likely to overcome short-term, material benefits that individual participants are tempted to pursue. In a different context – a large, heterogeneous group with no communication and no information about past trustworthiness who jointly use a common-pool resource – individuals will tend to pursue short-term material benefits and potentially destroy the resource.

Thus, using a broader theory of human behavior that includes the possibility that participants use reciprocity and cooperate in social dilemmas when they trust others will do the same, enables scholars to generate testable hypotheses based on combinations of structural variables as they interact to increase or decrease the likelihood of cooperation and net benefits occurring [see Weber, Kopelman, and Messick 2004 for a similar effort]. It is not possible, however, to link all of the structural variables identified above in a one definitive causal model given the large number of variables and that many of them depend for their impact on the value of other variables. For now, it is possible to illustrate this general approach with the framework shown in Figure 3 where the structural variables discussed above are linked in a general way to the core relationships.

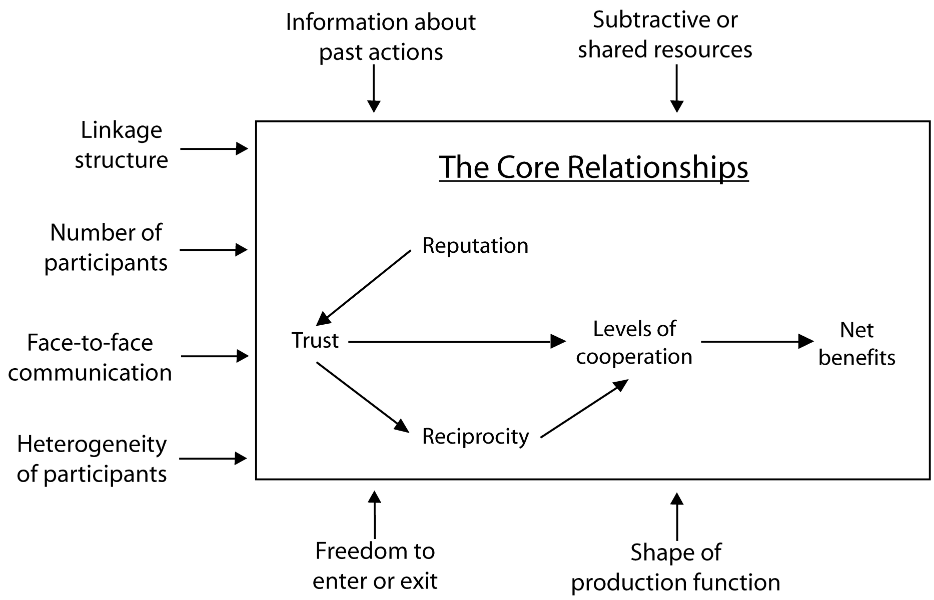

FIG. 3. A framework linking structural variables to the core relationships in a focal dilemma area.

One cannot assign a fixed direction of relationships in this approach, however, since the sign depends on the configuration of other variables in a particular focal social dilemma. A small group with extreme heterogeneity in the benefits to be obtained from a collective action, for example, is an entirely different group than a small group of relatively homogeneous players. Further, in a small group with extreme heterogeneity, face-to-face communication may lead to exacerbated conflict rather than reduction in conflict and agreement on new sets of rules. Instead of one large, general causal model, one can develop specific scenarios of causal direction, such as those posited above, that can be tested. Thus, an important next step in the development of collective-action theory is more careful attention to how structural variables interact with one another. One cannot posit simple explanations based upon an assumption that size alone makes a difference, that heterogeneity alone makes a difference, that a step level production function alone makes a difference, or the capacity to exit alone makes a difference – all proposed by some scholars as the primary variable one needs to examine. It is the combination of these variables that evoke norms, help or hinder building reputations and trust, and enable effective or destructive interactions and learning to occur. What is important about this simple and general framework is recognition that at any one time multiple variables affect the core variables of reputation, trust, and reciprocity.

Further, the variables linked together on Figure 3 are not an exhaustive set of all structural variables posited to affect collective action – they are the set that appears to be most frequently mentioned in the general literature reviewed above. Still other variables are identified in more specialized work. Agrawal [2002] has, for example, identified over 30 variables posited by scholars studying collective action related to organizing the governance of common-pool resources. Many of the variables he identifies have interactional effects. Agrawal [*ibidem*: 68-70] develops several causal chains to connect a subset of these variables together for testing in field and laboratory settings. Some of the variables identified by Agrawal relate to the likelihood of participants changing the rules that affect the structural variables that, in turn, affect the core relationships. Given the importance of changing rules to solve collective-action problems, let us now turn to a brief discussion of changing rules.

Changing Rules

Endogenous institutional change has not been as thoroughly explored as other ways of coping with social dilemmas because there is no reason to think that participants may change rules when using the simple models of rational choice. In the earlier work on collective action, theorists presumed that external policymakers would need to change the rules so as to impose positive or negative incentives on participants. Now that more theorists take the possibility of a broader theory of human behavior into account, theoretical work on the multiple levels of collective action involved in changing rules is highly relevant.

Designing rules to cope effectively with social dilemma situations is a challenging problem whether undertaken by participants or external actors. Individuals must invest time and energy in contemplating alternative rules and how they can be monitored and enforced. Individuals who are engaged in a repetitive social dilemma may find it worthwhile it to invest time and effort in coming up with new rules to restructure the situation in which they find themselves so as to make it easier to solve a collective-action problem over the long run.

Rules imply that a group of individuals have developed shared understandings that certain actions in particular situations must, must not, or may be undertaken and that sanctions will be taken against those who do not conform [Crawford and Ostrom 2005]. Rules tend to be self-conscious artifacts related to particular actions in specific situations [V. Ostrom 1999]. Rules are created in both private associations as well as in more formalized public institutions where they carry the additional legal weight of being enforced legal enactments [Williamson 2000].

Rules may be used to affect the size of group or its heterogeneity by making contributions mandatory, more obvious, or proportional to the benefits that are received. Rules that enable clear-cut ways of reducing the risk of being a sucker and contributing resources when others do not are an important technique for improving outcomes. The analysis of specific rules that might be used to solve the problems of collective action yields an immense variety of rules [see E. Ostrom 2005]. Many of the rules that are crafted by the individuals themselves do seek to account for the specific combination of structural variables that they face rather than simply creating a blueprint set of rules recommended for some general scenario.

To change rules those involved in collective action must shift out of a current "game" to a deeper level arena. All rules are the result of decisions made in a deeper arena that define how rules may be changed. The nesting of rules within rules at several levels is similar to the nesting of computer languages. What can be done at one level depends on the capabilities and limits of the rules (or the software) at that level or at other levels. Changes in the rules used to improve outcomes at one level occur within a currently "fixed" set of rules at an even deeper level. Changes in deeper-level rules usually are more difficult and more costly to accomplish, thus increasing the stability of mutual expectations among individuals interacting according to a set of rules.[4] In Figure 4, for simplicity sake, we nest the structural variables of a focal social dilemma in a set of rules that are made in a single, linked rule-making arena.

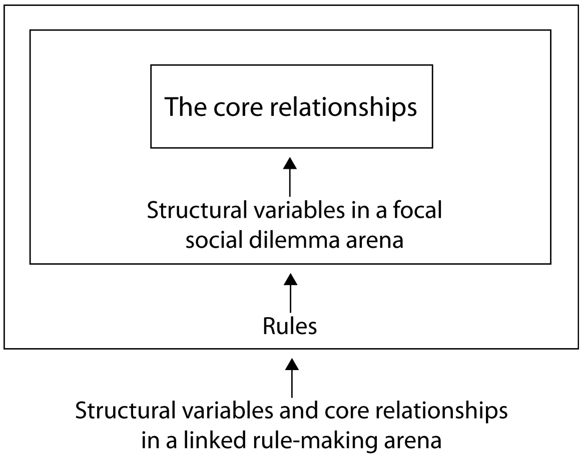

FIG. 4. Rulers affecting the structural variables of a focal social dilemma arena decided upon in a linked rule-making arena.

The capacity to change rules in an effort to improve net individual and group benefits depends on a large number of variables, including the type of rules affecting the structure of the deeper rule-changing arena [Knight 1992; E. Ostrom 2005]. The rules of the deeper level may assign differential advantages to participants in the rule changing process. Those with the least voice in collective choice processes may lose as a result of rule changes even though the aggregate benefit is greater. Whether endogenous change occurs in the rules and the distribution of costs and benefits depends on the distribution of assets among participants, the rules used to change rules, and the level of exogenous pressure on participants.

[4] It is useful to distinguish three levels of rules that cumulatively affect the actions taken and outcomes obtained in any setting. Sproule-Jones [2002: 70] calls these "rule stacks":

1. *Operational rules* affect day-to-day decisions made by participants because of their impact on the structural variables of that setting.
2. *Collective-choice rules* affect operational activities and outcomes through their effects in determining who is eligible to make operational rules and the specific collective-choice rules to be used in changing operational rules.
3. *Constitutional-choice rules* affect operational activities and outcomes in determining who is eligible and the rules to be used in crafting the set of collective-choice rules that in turn affect the set of operational rules.

At each level of analysis there may be one or more arenas in which the types of decisions made at that level will occur. Policy making regarding the rules that will be used to regulate operational-level action situations is usually carried out in one or more collective-choice arenas as well as being enforced at an operational level. Dilemmas are not limited to an operational level of analysis. They frequently occur at the collective-choice and constitutional levels of analysis.

The impact of a change in rules is likely to impact differentially on participants [Knight 1992]. Among the costs involved in changing rules are the time and resources needed to 1) compare the expected net returns of alternative rule changes, 2) make decisions about future rules in the relevant collective choice arena linked to any particular operational situation (which increase if rule changes differentially impact on participants), 3) explain new rules to those involved and gain their acceptance, 4) monitor and sanction rule infractions, and 5) resolve conflicts over rule interpretations and enforcement strategies [E. Ostrom 1990, 1999].

In many situations, the positive returns of solving social dilemmas are substantial enough that strong incentives exist to invest in the process of establishing a better set of rules to structure these future sets of interactions. The possibility of error in estimating the positive returns that could be achieved and/or the costs of any of the above steps is also considerable – particularly in complex situations that are strongly affected by external and unpredictable factors. Further, strategic actors may try to change rules in such a manner that others bear the costs of the change while those initiating the change obtain most of the benefits.

Rules designed by external authorities can also improve or reduce the value of the outcomes achieved. The adoption of non-optimal rules by officials external to particular situations can occur due to lack of knowledge regarding the local situation or perverse incentives facing the officials. Research on hierarchies based on principle-agent models has also generated a fuller recognition that these mechanisms are themselves replete with many social dilemma situations that are not conducive to the creation of optimal rules by public officials [Araral 2005]. Thus, no bloodless institutional designer exists inside the black box of "the state" that searches out the optimal set of rules for a situation and engages in perfect monitoring and enforcement of these rules. Simply calling on an abstract, reified entity – "the state" – and assuming that it can change the rules affecting a social dilemma is not an adequate solution to any collective-action problem [Lichbach 1996; Brennan and Pettit 2004].

No doubt exists about the importance of governments – at supra-national, national, and subnational levels – in potentially coping with problems of collective action. It is particularly important, however, also to explore how a wide diversity of institutions that are neither markets nor states operates in diverse field settings to enhance the joint benefits that individuals achieve in collective-action situations. Many of these institutions are constituted by participants in a self-governing process rather than being imposed by external authorities [V. Ostrom 1997].

Conclusions

A key lesson of research on collective-action is recognizing the complex linkages among variables at multiple levels that together affect individual reputations, trust and reciprocity as these, in turn, affect levels of cooperation and joint benefits. Conducting empirical research on collective action is thus extremely challenging. There is no way that one can analyze the entire “spaghetti plate” of variables that have been identified and their interactions in a single empirical analysis. The reason that experimental research has become such an important method for testing theory is that it is a method for controlling the setting of many variables while changing only one or two variables at a time [Camerer 2003]. In addition, one can self-consciously examine the interaction of several variables over a series of carefully designed experiments – something that is almost impossible to do in field research. Conducting research in similar environments that differ in regard to one or two key variables is also an important strategy, but very difficult to find such settings. Large-N research on collective action is a challenge both in terms of obtaining accurate and consistent data, but also because of the large number of variables that potentially affect any one type of collective action [Poteete and Ostrom 2004]. Instead of looking at all of the potential variables, one needs to focus in on a well defined but narrow chain of relationships – as recommended by Agrawal [2002]. One can then conduct analysis of a limited set of variables that are posited to have a strong causal relationship [for examples, see Gibson, Williams and Ostrom 2005; Hayes and Ostrom 2005].

An earlier version of this paper was prepared for the Oxford Handbook of Comparative Politics, ed. Carles Boix and Susan Stokes. Oxford, UK: Oxford University Press. Presentations of earlier drafts were made at Duke University, Durham, NC, April 28, 2005, and at the Workshop in Political Theory and Policy Analysis, Indiana University, Bloomington, May 5, 2005. I acknowledge with deep appreciation the extensive comments of Arun Agrawal, Geoffrey Brennan, Michael McGinnis, Lesa Morrison, Roger Parks, Michael Schoon, Suzanne Shanahan, and Lihua Yang. The excellent editing of Patty Lezotte and David Price has been of great help.

References

Abreau, D.

1988 “On the theory of infinitely repeated games with discounting.” *Econometrica* 80 (4): 383-396.

Agrawal, A.

2000 "Small is beautiful, but is larger better? Forest-management institutions in the Kumaon Himalaya, India." Pp. 57-86 in *People and Forests: Communities, Institutions, and Governance*, edited by C. Gibson, M. McKea, and E. Ostrom. Cambridge, MA: MIT Press.

2002 "Common resources and institutional sustainability." Pp. 41-85 in *The Drama of the Commons. National Research Council, Committee on the Human Dimensions of Global Change*, E. Ostrom, T. Dietz, N. Dolšak, P.C. Stern, S. Stonich, and E. Weber. Washington, DC: National Academy Press.

Ahn, T.K., Janssen, M., and Ostrom, E.

2004 "Signals, symbols, and human cooperation." Pp. 122-139 in *The Origins and Nature of Sociality*, edited by R.W. Sussman and A.R. Chapman. New York: de Gruyter.

Alchian, A.A.

1950 "Uncertainty, evolution, and economic theory." *Journal of Political Economy* 58 (3): 211-221.

Andreoni, J.

1989 "Giving with impure altruism: Applications to charity and Ricardian equivalence." *Journal of Political Economy* 97: 1447-1458.

Araral, E.

2005 "Bureaucratic incentives, path dependence, and foreign aid: An empirical institutional analysis of irrigation in the Philippines." *Policy Sciences* 38 (2-3): 131-157.

Axelrod, R.

1984 *The Evolution of Cooperation*. New York: Basic Books.

Axelrod, R., and Hamilton W.D.

1981 "The evolution of cooperation." *Science* 211: 1390-1396.

Bardhan, P.

1993 "Analytics of the institutions of informal cooperation in rural development." *World Development* 21 (4): 633-639.

Bates, R.H., and Shepsle, K.A.

1995 "Demographics and institutions." Paper presented at the Frontiers of Economics Conference (in honor of Douglass C. North), Washington University, St. Louis, Missouri, March 17-19.

Bendor, J., and Mookherjee, D.

1987 "Institutional structure and the logic of ongoing collective action." *American Political Science Review* 81 (1): 129-154.

Bolton, G.E., and Ockenfels, A.

2000 "ERC: A theory of equity, reciprocity, and competition." *American Economic Review* 90: 166-193.

Boyd, R., and Richerson, P.J.

1988 "The evolution of reciprocity in sizable groups." *Journal of Theoretical Biology* 132: 337-356.

1992 "Punishment allows the evolution of cooperation (or anything else) in sizable groups." *Ethology and Sociobiology* 13: 171-195.

Brennan, G., and Pettit, P.

2004 *The Economy of Esteem*. Oxford: Oxford University Press.

Camerer, C.F.

2003 *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Cameron, L.A.

1999 "Raising the stakes in the ultimatum game: Experimental evidence from Indonesia." *Economic Inquiry* 37 (21): 47-59.

Cardenas, J.-C.

2000 "How do groups solve local commons dilemmas? Lessons from experimental economics in the field." *Environment, Development and Sustainability* 2: 305-322.

Carpenter, J., Burks, S., and Verhoogen, E.

2005 "Comparing students to workers: The effects of social framing on behavior in distribution games." In *Field Experiments in Economics*, edited by J. Carpenter, G. Harrison, and J.A. List. Greenwich, CT: JAI Press.

Chamberlin, J.

1974 "Provision of collective goods as a function of group size." *American Political Science Review* 68: 707-716.

Cook, K.S., and Hardin, R.

2001 "Norms of cooperativeness and networks of trust." Pp. 327-357 in *Social Norms*, edited by M. Hechter and K.-D. Opp. New York: Russell Sage Foundation.

Cox, J.C.

2004 "How to identify trust and reciprocity." *Games and Economic Behavior*. 46: 260-81.

Cox, J.C., and Deck, C.

2005 "On the nature of reciprocal motives." *Economic Inquiry* 43 (3): 623-635.

Cox, J.C., Friedman, D., and Gjerstad, S.

Forth "A tractable model of reciprocity and fairness." *Games and Economic Behavior.*

Crawford, S.E.S., and Ostrom, E.

2005 "A grammar of institutions." Pp. 137-174 in *Understanding Institutional Diversity*, edited by E. Ostrom. Princeton, NJ: Princeton University Press. Originally published in *American Political Science Review* 89 (3): 582-600, 1995.

Davis, D.D., and Holt, C.A.

1993 *Experimental Economics*. Princeton, NJ: Princeton University Press.

Dawes, R.M., Orbell, J.M., and van de Kragt, A.

1986 "Organizing groups for collective action." *American Political Science Review* 80 (4): 1171-1185.

Dietz, T., Ostrom, E., and Stern, P.

2003 "The struggle to govern the commons." *Science* 302 (5652): 1907-1912.

Falk, A., Fehr, E., and Fischbacher, U.

2002 "Appropriating the commons: A theoretical explanation." Pp. 157-192 in *The Drama of the Commons. National Research Council, Committee on the Human Dimensions of*

Global Change, edited by E. Ostrom, T. Dietz, N. Dolšak, P.C. Stern, S. Stonich, and E. Weber. Washington, DC: National Academy Press.

Fiske, S.T., Harris, L.T., and Cuddy, A.J.C.

2004 “Why ordinary people torture enemy prisoners.” *Science* 306: 1482-1483.

Frohlich, N., and Oppenheimer, J.

1970 “I get by with a little help from my friends”. *World Politics*. 23: 104-20.

1998 “Some consequences of e-mail vs. face-to-face communication in experiment.” *Journal of Economic Behavior and Organization* 35: 389-403.

Fudenberg, D., and Maskin, E.

1986 “The folk theorem in repeated games with discounting or with incomplete information.” *Econometrica* 54 (3): 533-554.

Gellar, S.

2005 *Democracy in Senegal: Tocquevillian Analytics in Africa*. New York: Palgrave Macmillan.

Gibson, C., Williams, J., and Ostrom, E.

2005 “Local enforcement and better forests.” *World Development* 33 (2): 273-284.

Gigerenzer, G., and Selten, R. (eds.)

2001 *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press.

Granovetter, M.

1973 “The strength of weak ties.” *American Journal of Sociology* 78: 1360-1380.

Güth, W., and Kliemt, H.

1995 “Competition or co-operation: On the evolutionary economics of trust, exploitation and moral attitudes.” Working Paper. Berlin, Germany: Humboldt University.

Hardin, R.

1976 “Group provision of step goods.” *Behavioral Science* 21: 101-106.

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

Hauert, C., De Monte, S., Hofbauer, J., and Sigmund, K.

2002 “Volunteering as red queen mechanisms for cooperation in public good games.” *Science* 296: 1129-1132.

Hauk, E., and Nagel, R.

2001 “Choice of partners in multiple two-person prisoner’s dilemma games: An experimental study.” *Journal of Conflict Resolution* 45: 770-793.

Hayes, T.M., and Ostrom, E.

2005 “Conserving the world’s forests: Are protected areas the only way?” *Indiana Law Review* 37 (3): 595-617.

Heckathorn, D.D.

1991 “Collective sanctions and group heterogeneity: Cohesion and polarization in normative systems.” Working Paper. Storrs: University of Connecticut, Department of Sociology.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., and Gintis, H. (eds)

2004 *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press.

Isaac, R.M., Mathieu, D. and Zajac, E.E.

1991 "Institutional framing and perceptions of fairness." *Constitutional Political Economy* 2 (3): 329-370.

Isaac, R.M., and Walker, J.M.

1988 "Communication and free-riding behavior: The voluntary contribution mechanism." *Economic Inquiry* 26 (4): 585-608.

Janssen, M.A.

2004 "Evolution of cooperation when feedback to reputation scores is voluntary". Working Paper. Bloomington: Indiana University, Center for the Study of Institutions, Population, and Environmental Change.

2005 "Evolution of cooperation in a one-shot prisoner's dilemma based on recognition of trustworthy and untrustworthy agents". Working Paper. Bloomington: Indiana University, Center for the Study of Institutions, Population, and Environmental Change.

Johnson, R., and Libecap, G.

1982 "Contracting problems and regulation: The case of the fishery." *American Economic Review* 72: 1005-1023.

Jones, B.D.

2001 *Politics and the Architecture of Choice: Bounded Rationality and Governance*. Chicago: University of Chicago Press.

Jones, E.C.

2004 "Wealth based trust and the development of collective action." *World Development* 32 (4): 691-711.

Keohane, R.O.

1984 *After Hegemony*. Princeton, NJ: Princeton University Press.

Kerr, N.L., and Kaufman-Gilliland, C.M.

1994 "Communication, commitment and cooperation in social dilemmas." *Journal of Personality and Social Psychology* 66: 513-529.

Knack, S.

1992 "Civic norms, social sanctions, and voter turnout." *Rationality and Society* 4: 133-156.

Knight, J.

1992 *Institutions and Social Conflict.* Cambridge: Cambridge University Press.

Kreps, D., Milgrom, P., Roberts, J., and Wilson, R.

1982 "Rational cooperation in finitely repeated prisoners' dilemma." *Quarterly Journal of Economics* 103: 179-200.

Libecap, G., and Wiggins, S.

1984 "Contractual responses to the common pool: Prorationing of crude oil production." *American Economic Review* 74: 87-98.

Lichbach, M.I.

1995 *The Rebel's Dilemma*. Ann Arbor: University of Michigan Press.

1996 *The Cooperator's Dilemma.* Ann Arbor: University of Michigan Press.

Luce, R.D., and Raiffa, H.

1957 *Games and Decisions: Introduction and Critical Survey*. New York: Wiley.

Marwell, G., and Oliver, P.

1993 *The Critical Mass in Collective Action: A Micro-Social Theory*. New York: Cambridge University Press.

Morikawa, T., Orbell, J.M., and Runde, A.S.

1995 "The advantage of being moderately cooperative." *American Political Science Review* 89 (3): 601-611.

Olson, M.

1965 *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.

Orbell, J.M., and Dawes, R.M.

1991 "A 'cognitive miser' theory of cooperators' advantage." *American Political Science Review* 85 (2): 515-528.

Orbell, J.M., Morikawa, T., Hartwig, J., Hanley, J., and Allen, N.

2004 "Machiavellian Intelligence as a Basis for the Evolution of Cooperative Dispositions." *American Political Science Review* 98 (1): 1-15.

Ostrom, E.

1990 *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.

1998 "A behavioral approach to the rational choice theory of collective action." *American Political Science Review* 92 (1): 1-22.

1999 "Coping with the tragedies of the commons." *Annual Review of Political Science* 2: 493-535.

2001 "Social dilemmas and human behavior." Pp. 21-41 in *Economics in Nature: Social Dilemmas, Mate Choice and Biological Markets*, edited by R. Noë, J. Van Hooff, and P. Hammerstein. Cambridge: Cambridge University Press, 21-41.

2005 *Understanding Institutional Diversity*. Princeton, NJ: Princeton University Press.

Ostrom, E., Gardner, R., and Walker, J.

1994 *Rules, Games, and Common-Pool Resources*. Ann Arbor: University of Michigan Press.

Ostrom, V.

1999 "Taking constitutions seriously: Buchanan's challenge to twentieth-century political science." Pp. 123-136 in *Competition and Cooperation: Conversations with Nobelists about Economics and Political Science* edited by J.E. Alt, M. Levi, and E. Ostrom. New York: Russell Sage Foundation.

1997 *The Meaning of Democracy and the Vulnerability of Democracies: A Response to Tocqueville's Challenge.* Ann Arbor: University of Michigan Press.

Ostrom, V., and Ostrom, E.

1999 "Public goods and public choices." Pp. 107-118 in *Polycentricity and Local Public Economies: Readings from the Workshop in Political Theory and Policy Analysis*, edited by M. McGinnis. Ann Arbor: University of Michigan Press, 107-18.

Panchanathan, K., and Boyd, R.

2004 "Indirect reciprocity can stabilize cooperation without the second-order free rider problem." *Nature* 432: 499-502.

Parks, C.D., Henager, R.F., and Scamahorn, S.D.

1996 "Trust and reactions to messages of intent in social dilemmas." *Journal of Conflict Resolution* 40: 134-151.

Pecorino, P.

1999 "The effect of group size on public good provision in a repeated game setting." *Journal of Public Economics* 72: 121-134.

Poteete, A., and Ostrom, E.

2004 "Heterogeneity, group size and collective action: the role of institutions in forest management." *Development and Change* 35 (3): 437-461.

Rapoport, A.

1997 "Order of play in strategically equivalent games in extensive form." *International Journal of Game Theory* 26 (1): 113-36.

Richerson, P.J., and Boyd, R.

2005 *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago: University of Chicago Press.

Sally, D.

1995 "Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958-1992." *Rationality and Society* 7: 58-92.

Satz, D., and Ferejohn, J.

1994 "Rational choice and social theory." *Journal of Philosophy* 91 (2): 71-87.

Schelling, T.C.

1978 *Micromotives & Macrobehavior.* New York: W.W. Norton.

Seabright, P.

1993 "Managing local commons: Theoretical issues in incentive design." *Journal of Economic Perspectives* 7 (4): 113-134.

Sen, A.K.

1977 "Rational fools: A critique of the behavioral foundations of economic theory." *Philosophy and Public Affairs* 6: 317-344.

Shivakumar, S.J.

2005 *The Constitution of Development: Crafting Capabilities for Self-Governance*. New York: Palgrave Macmillan.

Simon, H.A.

1955 "A behavioural model of rational choice." *Quarterly Journal of Economics* 69: 99-188.

1999 "The potlatch between political science and economics." Pp. 112-119 in *Competition and Cooperation: Conversations with Nobelists about Economics and Political Science*, edited by J. Alt, M. Levi, and E. Ostrom. New York: Russell Sage Foundation.

Sproule-Jones, M.

2002 *Restoration of the Great Lakes: Promises, Practices, Performances*. Vancouver: UBC Press.

Valley, K.L., Moag, J., and Bazerman, M.H.

1998 "A matter of trust: effects of communication on the efficiency and distribution of outcomes." *Journal of Economic Behavior and Organization* 34 (2): 211-238.

van de Kragt, A., Orbell, J.M., and Dawes, R.M.
1983 "The minimal contributing set as a solution to public goods problems." *American Political Science Review* 77 (1): 112-122.

Weber, J.M., Kopelman, S., and Messick, D.M.
2004 "A conceptual review of decision making in social dilemmas: Applying a logic of appropriateness." *Personality and Social Psychology Review* 8: 281-307.

Weissing, F.J., and Ostrom, E.
1991 "Irrigation institutions and the games irrigators play: Rule enforcement without guards." Pp. 188-262 in *Game Equilibrium Models II: Methods, Morals, and Markets*, edited by R. Selten. Berlin: Springer-Verlag.

Wiggins, S., and Libecap, G.
1987 "Firm heterogeneities and cartelization efforts in domestic crude oil." *Journal of Law, Economics, and Organization* 3: 1-25.

Williamson, O.E.
2000 "The new institutional economics: Take stock, looking ahead." *Journal of Economic Literature* 38: 593-613.

Yamagishi, T., and Cook, K.S.
1993 "Generalized exchange and social dilemmas." *Social Psychology Quarterly* 56 (4): 235-248.

Collective Action and Local Development Processes

Abstract: Developing a coherent theory of collective action that is also relevant for practice in explaining local development is a major challenge. At the individual level, individuals do take costly actions that effectively take the interests of others into account in many field and experimental settings but this is not consistent with contemporary game theory. We need to move ahead to achieve a more coherent synthesis of theoretical work that posit variables affecting the success or failure diverse forms of collective action. The first section of this paper discusses the growing and extensive theoretical literature positing a large number of structural variables presumed to affect the likelihood of individuals achieving collective action to overcome social dilemmas. None of these structural variables, however, would change predictions if one uses the *model* of rationality that has proved successful in explaining behavior and outcomes in competitive market settings as a universal *theory* of human behavior. Thus, the second section examines how a theory of boundedly rational, norm-based human behavior is a better foundation for explaining collective action than a model of maximizing material payoffs to self. The third section examines the linkage between the structural measures first discussed with the individual relationships discussed in the second. The fourth section looks at how changing the rules of a focal dilemma in deeper arenas in efforts to improve the net benefits from collective action by affecting the structural variables of the focal arena. The conclusion reflects on the challenge that social scientists face in testing collective-action theory in light of the large number of variables posited to affect outcomes.

Keywords: collective action, social-ecological systems (SESs), inter-disciplinary research, multi-level development, sustainability.

Elinor Ostrom is Arthur F. Bentley Professor of Political Science; Co-Director of the Workshop in Political Theory and Policy Analysis, Indiana University, Bloomington; and Founding Director, Center for the Study of Institutional Diversity, Arizona State University. She is a member of the National Academy of Sciences, the American Academy of Arts and Sciences, and a recipient of the Frank E. Seidman Prize in Political Economy and the Johan Skytte Prize in Political Science. Her books include *Governing the Commons*; *Rules, Games, and Common-Pool Resources* (with Roy Gardner and James Walker); *Local Commons and Global Interdependence* (with Robert Keohane); *The Commons in the New Millennium* (with Nives Dolšak); *The Samaritan's Dilemma: The Political Economy of Development Aid* (with Clark Gibson, Krister Andersson, and Sujai Shivakumar); and *Understanding Institutional Diversity*.