Journal of Economic Behavior & Organization
Vol. 61 (2006) 149–163

JOURNAL OF Economic Behavior & Organization

www.elsevier.com/locate/econbase

The value-added of laboratory experiments for the study of institutions and common-pool resources

Elinor Ostrom *

Workshop in Political Theory and Policy Analysis, Department of Political Science, Indiana University, 513 North Park, Bloomington, IN 47408-3895, USA

Received 7 June 2004; received in revised form 28 January 2005; accepted 23 February 2005
Available online 12 May 2006

Abstract

This article provides an overview of the effects of communication on experimental studies of behavior and outcomes in common-pool resource (CPR) dilemmas. Aggregate outcomes in CPR dilemmas without communication approximate predictions of non-cooperative game theory, but allowing cheap talk results in higher outcomes. When exogenous rules are monitored at realistic levels, subjects cheat even though following the rule would generate optimal outcomes. If given the opportunity, experimental subjects will devise their own rule systems and impose sanctions on each other. These findings complement field research on more complex resources and communities by confirming the critical importance of communication and endogenous rule formation.

JEL classification: B52; C72; C92

Keywords: Institutions; Common-pool resources; Experimental findings

1. Introduction

Given my reputation as an avid field researcher, colleagues often ask why I 'bother' with conducting experiments. They ask questions such as 'Why would you pay any attention to outcomes in an experiment?' and 'What more can you possibly learn about institutions and resource governance from laboratory experiments that you have not already learned in the field?'

My first response is that we should learn more from multiple research methods applied to the same question than from a single method. Further, experimental research enables one to test the

* Tel.: +1 812 855 0441; fax: +1 812 855 3150.
E-mail address: ostrom@indiana.edu.

0167-2681/$ – see front matter
doi:10.1016/j.jebo.2005.02.008

impact of specific variables in repeated controlled settings, something that is never available to a scholar studying field settings. One gains external validity in doing field research and internal validity in the laboratory. When a researcher can use both methods related to one theoretical set of questions, the scientific community can have more confidence in the results.

In this article, I would like to provide a summary of what we have learned from experimental research about the governance of common-pool resources. Fieldwork teaches us that many biophysical and social factors affect the likelihood of resource users overcoming perverse incentives to organize themselves, establish boundaries, craft rules for sustainable harvesting, monitor their own rules, and sanction colleagues who break these rules. In the field, we do *not* observe universal success in overcoming the multiple dilemmas that resource users face. In many settings, however, users have behaved contrary to the theoretical predictions derived in resource economics textbooks of the 1980s and 1990s that managing common-pool resources requires either government-imposed management or the division of a resource into private property.[1]

When conducting field research, one of the frustrating aspects is that so many variables are involved that one is never certain that one has isolated the specific variable (or limited set of variables) that causes an outcome. A good way to understand which components of a common-pool resource situation affect behavior and outcomes, and how, is to study a simplified version in an experimental laboratory. In the laboratory, the researcher carefully establishes the specific components of the theoretical situation to be studied and controls other variables so they do not confound the analysis.

Thus, in this article, I will provide a general overview of the experimental studies of behavior and outcomes in common-pool resource dilemmas conducted by Roy Gardner, James Walker, and myself; by other colleagues at Indiana University; or by colleagues at other universities who have replicated key experimental designs. Given the space limitations of this article, I will refer the reader for details of specific designs and data from them to published papers or books. In this article, I focus primarily on the lessons we have learned from these experiments as a core part of a larger research program combining theoretical developments, field research, and laboratory experiments.

In Section 2, a common-pool resource and an appropriation dilemma are defined. Section 3 describes the sparse design we developed for our baseline experiment to include just the core aspects that underlie appropriation dilemmas in field settings: a non-linear payoff function based on the decisions of more than two appropriators who received no information about each others' actions. To keep the baseline as simple as possible, the subjects were homogeneous in their endowments and were not allowed to communicate. The average payoffs achieved in the baseline experiments did approach the Nash equilibrium of the non-cooperative game, but with some unexplained pulsing behavior which was replicated by agent-based models created by Deadman (1997) and by Jager and Janssen (2002).

In Section 4, experiments that introduced face-to-face communication are described. In these experiments, subjects were authorized to communicate with one another in a group setting before returning to their terminals to make their own private decisions. Introducing an opportunity for 'cheap talk' in a dilemma setting where agreements are not enforced by an external authority is

[1] Our experiments have been stimulated by extensive field research on the impact of institutions on the performance of irrigation systems (Ostrom, 1992; Tang, 1992; Lam, 1998; Shivakoti and Ostrom, 2002), inshore fisheries (Schlager, 1994; Acheson, 2003; Acheson and Gardner, 2004; Yandle and Dewees, 2003), pastoral groups (Agrawal, 1999; Kaul, 1996), and forest governance (Gibson et al., 2000, 2005; Varughese and Ostrom, 2001; Poteete and Ostrom, 2004; Moran and Ostrom, 2005).

viewed within the context of non-cooperative game theory as irrelevant. The same outcome is predicted as in the baseline experiment. The experimental findings were, however, dramatically different. An initial set of communication experiments retained the homogeneity of asset structure but examined the difference between allocating subjects a small asset endowment versus a larger endowment. Rocco and Warglien (1995) replicated the findings from our no-communication and face-to-face communication experiments but found that communication via electronic means was not as effective. The last part of this section provides an overview of the experiments exploring whether communication would help to overcome the problems of heterogeneity of asset structure.

Section 5 provides an overview of the series of experiments where we changed the payoff component to allow subjects to sanction one another at a cost to themselves. Since using this option produces a benefit for all at a cost to the individual, the game-theoretic prediction is that no one will choose the costly sanctioning option. We had observed frequent instances of users sanctioning one another in the field and wanted to explore whether this would occur in a carefully designed environment controlling for other potentially confounding variables. Our results established the willingness of subjects to punish each other for actions inconsistent with their promises to one another.

Section 6 syntheses the experiments where we changed the authority rule to allow subjects to covenant with one another to determine their investment levels and to adopt a sanctioning system if they wished. Again, the predicted outcome is the same as the baseline experiment since the subjects now faced a multi-level dilemma and had to invest their own resources to enforce their decisions. In all three of these structurally changed appropriation experiments, however, subjects demonstrate their willingness and ability to search out and adopt better outcomes than those predicted. Section 7 reviews experiments where rules were exogenously imposed. Section 8 is a conclusion. Let us now turn to the concept of a common-pool resource.

2. The definition of a common-pool resource

A common-pool resource such as an irrigation system, a fishing ground, a forest, the Internet, or the stratosphere is a natural or man-made resource from which it is difficult to exclude or limit users once the resource is provided by nature or humans (Ostrom et al., 1994). One person's consumption of resource units, such as water, removes those units from the resource system. Thus, the water or fish harvested by one user are no longer available for others. The difficulty of excluding beneficiaries is a characteristic that is shared with public goods while the subtractability of the resource units is shared with private goods.

When the resource units produced by a common-pool resource are highly valued and institutional constraints do not restrict harvesting (an open-access situation), individuals face strong incentives to appropriate more and more resource units leading eventually to congestion, overuse, and even the destruction of the resource itself. If some individuals reduce their appropriation levels, the benefits are shared with others whether they restrict their harvesting or not. Some individuals may free ride on the costly actions of others.

Consequently, the joint users of a common-pool resource face a dilemma given the incentive for individuals to appropriate (harvest) more resource units when acting independently than they would if they could find some way of coordinating their appropriation activities. The characteristics of each of these problems differ substantially. In this review, I focus on static appropriation dilemmas since they are what most readers associate with the well-known 'tragedy of the commons' (Hardin, 1978; see Dietz et al., 2003).

3. The baseline appropriation experimental design

In our initial effort to design a baseline experiment, we wanted to start with an experiment that was as simple as we could specify without losing crucial aspects of the problems that appropriators face in the field. To approximate some of the complexity of field settings (as well as the classic textbook example of a common-pool resource), however, we wanted to examine behavior in an appropriation situation with a non-linear transformation function based on Gordon (1954) and Scott (1955). We also wanted to include a sufficient number of players whose knowledge of outcomes did not automatically provide information about each player's actions, again similar to the field.

The first question we wished to address with a baseline experiment was whether subjects' behavior in this stark dilemma situation would be similar to the Nash equilibrium predicted by non-cooperative game theory. The experiments were conducted at Indiana University by recruiting students from the general undergraduate population and indicating that they would be making decisions in an 'economic choice situation' where they earn funds depending on their own and others' decisions (see Ostrom et al., 1994 for fuller information about these experiments).

Subjects were assigned an equal number of tokens in each round and told they could invest in two markets. The first market yielded a fixed rate of return for each token invested. The second market (the CPR) yielded a rate of return dependent upon the total number of tokens invested by all subjects in an experiment. The payoffs assigned to tokens invested in the second market were a concave function, F that depended on the number of tokens, x_i, allocated to this market. Initially, the sum of the subjects' actions, $\sum x_i$, generated better outcomes than the safe fixed rate of return. If the subjects allocated a sufficiently large number of their tokens, the outcome they received was less than their best alternative. In other words, allocating too many assets to the common-pool resource would be counterproductive. Specifically, the payoff to a subject was given by

$$we, \quad \text{if } x_i = 0, \qquad w(e - x_i) + \left(x_i/\sum x_i\right) F\left(\sum x_i\right), \quad \text{if } x_i > 0. \tag{1}$$

If subjects put all of their tokens into Market 1, they received a certain return equal to the amount of their endowment times an unchanging rate, w. If subjects put some of their endowed assets into the outside option, Market 1, and the rest in Market 2 (the CPR), they received part of their return from the outside option and the rest from their proportional investment in the common-pool resource times the total output of the common-pool resource as determined by function F.

3.1. Predicted outcomes for a common-pool resource in the laboratory

In the baseline experiments, we utilized the following equation for the transformation function, F, measured in units of output (outcome units):

$$23\left(\sum x_i\right) - 25\left(\sum x_i\right)^2. \tag{2}$$

In the first set of experiments, we allocated 10 tokens to each subject in each round. In later experiments, we allocated 25 tokens. Their outside opportunity was valued at \$.05 per token. They earned \$.01 on each outcome unit they received from investing tokens in the common-pool resource. Subjects were informed that they would participate in an experiment that would last no more than two hours. The number of rounds in each experiment varied between 20 and 30 rounds. In addition to being told the payoff function specifically, subjects were provided with look-up tables that eased their task of determining outcomes depending on their own and others' decisions.

The predicted outcome for a finitely repeated game where subjects are not discounting the future and each subject is assumed to be maximizing monetary returns is for each subject to invest 8 tokens in the common-pool resource for a total of 64 tokens. By design, the prediction is the same for both endowment levels. At this level of investment, they would each earn $.66 per round in the 10-token experiments and $.70 per round in the 25-token experiments (players were paid one-half of their computer returns in the 25-token experiments to keep the payoffs roughly similar). The players could, however, earn considerably more if the total number of tokens invested was 36 tokens (rather than 64 tokens) in the common-pool resource. This optimal level of investment would earn each subject $.91 per round in the 10-token experiment and $.83 per round in the 25-token experiment.

3.2. *Behavior in a sparse experimental common-pool resource*

Subjects interacting in baseline experiments substantially overinvested. Subjects in the 10-token experiments achieved, on average, 37 percent of the maximum earning from the common-pool experiment available to them while subjects in the 25-token experiments received −3 percent (Ostrom et al., 1994, p. 116). At the individual level, however, subjects rarely invested the predicted level of investment at equilibrium. Instead, all experiments provided evidence of a strong pulsing pattern in which individuals appeared to increase their investments in the common-pool resource until there was a strong reduction in yield, at which time they tended to reduce their investments leading to an increase in yields. The pattern was repeated over time. At an aggregate level, behavior approximated the predicted Nash equilibrium in the 10-token experiments, but was far worse than predicted in the early rounds of the 25-token experiment. It only began to approach the predicted Nash level in later rounds.

No game-theoretical explanation yet exists for the substantial difference between the 10-token and the 25-token experiments. Those who study the impact of harvesting technology in the field, however, will feel quite comfortable with the difference between the 10- and 25-token experiments. Large-scale trawlers and mechanized forest-harvesting equipment are frequently thought to generate much less sustainable harvesting patterns than smaller boats or handsaws. The 25-token environment did provide subjects with more assets to devote to appropriation similar to a trawler versus a rowboat in a fishery.

3.3. *Replication of the baseline experiments in agent-based models*

An extremely interesting follow-up study was undertaken by Deadman in which artificial agents were programmed to use heuristics similar to those described by subjects in their post-experiment questionnaire responses. The simulated environment exactly replicated the baseline experiments. Deadman modeled agents who continued to increase their harvesting levels until their payoffs sharply declined and then reduced them until payoffs rebounded. The artificial agents consistently produced the same kind of pulsing returns and the consistent difference between 10- and 25-token environments. Deadman (1997, pp. 175–176) describes his results:

> As in CPR experiments, the group performance for the simulation follows an oscillating pattern in which high performance leads to over investment in the CPR and the resultant

> drop in performance causes a reduction in group-wide investment in the CPR... Still more interesting is the observation that the simulations perform similarly to subjects in laboratory experiments in terms of average performance over time. At the ten token endowment, the simulations perform near the Nash equilibria over time. At the 25 token endowment, the simulations perform near zero percent of optimum over time.

Jager and Janssen (2002) also developed a multi-agent model using the consummate framework derived from social psychology (see also Jager et al., 2002). They thought that they could replicate the data from the baseline CPR experiments described above with an assumption that individual subjects differed in regard to their social value orientations (SVOs). In their first series of simulations, they were indeed able, as Deadman had done independently, to explain the aggregated pattern of appropriation behavior. When they analyzed individual level data, Jager and Janssen were not as successful as they had been in simulating aggregate outcomes. Jager and Janssen (2002) suggested that the cognitive processes appear to be important in behavior. Those with low aspiration levels may lock into a habitual response too soon. Further, those agents 'conforming to the Homo psychologius have a better performance than the Homo economicus in approximating the empirical data' (p. 98).

4. Introducing communication

At an aggregate level, the findings from the baseline experiments were broadly consistent with predictions from non-cooperative game theory and not consistent with the findings from the field. An initial challenge was which component of the baseline experiment to change first in our effort to try to identify what factors in the field were more likely than others to increase levels of cooperation. An obvious factor in relatively small groups of appropriators is the capabilities of these groups to engage in face-to-face communication. Given that face-to-face communication had so often been shown to increase cooperation in other social dilemmas, we decided to focus first on this factor (see Ostrom and Walker, 1991).

4.1. Face-to-face communication among homogeneous subjects

In the repeated communication experiments, subjects first made 10 rounds of decisions in the context of the baseline appropriation game. After the 10th round, subjects listened to an announcement that told them they would have an open group discussion before each of the next rounds of the experiment. The subjects left their terminals and sat in a circle facing one another. After each discussion, they returned to their terminals to enter their anonymous decisions.

After each decision round, they learned what their aggregate investments had been, but not the decisions of individual players as in the baseline experiments. Thus, they learned whether total investments were greater than those they had agreed upon. While in many rounds, subjects did exactly as they had promised one another, some defections did occur over time. If promises were not kept, subjects used this information about the aggregate investment levels to castigate the unknown participant who had not kept to their agreement.

This opportunity for repeated face-to-face communication was extremely successful in increasing joint returns in the 10-token experiments where subjects obtained close to 100 percent of the maximum available returns. There were only 19 instances out of 368 total opportunities where a subject invested more in the common-pool resource than agreed upon for a 5 percent defection rate (Ostrom et al., 1994, p. 154). In the 25-token experiments, subjects also improved their overall performance. The temptation to defect, however, was greater in the 25-token experiments. Our conclusion in completing an analysis of these experiments was

> Communication discussions went well beyond discovering what investments would generate maximum yields. A striking aspect of the discussion rounds was how rapidly subjects, who had not had an opportunity to establish a well-defined community with strong internal norms, were able to devise their own agreements and verbal punishments for those who broke those agreements... In many cases, statements like 'some scumbucket is investing more than we agreed upon' were a sufficient reproach to change defectors' behavior. (Ostrom et al., 1994, p. 160; see also Simon and Gorgura, 2003)

The findings from these initial communication experiments are consistent with a large number of studies of the impact of face-to-face communication on the capacity of subjects to solve a variety of social dilemma problems (see Sally, 1995 for an extensive review).

4.2. Replications of face-to-face communication among homogeneous subjects

One of the great advantages of laboratory experiments is that they can be replicated as well as modified by other researchers so that one gains ever greater confidence in the findings. One of the first replications of our experiments was conducted by Rocco and Warglien, who found similar outcomes in the baseline, no-communication situation as well as in a replication of the face-to-face communication setting. They were interested, in addition, in the question of whether similar results could be obtained in a setting where communication was not organized on a direct face-to-face basis, but rather via a computerized exchange of information. Thus, they used identical structural variables as we had earlier for their own face-to-face design and then the same structural variables for a third design where they limited communication to a form of computerized exchange. The effect of communication, when not conducted on a face-to-face basis, was greatly reduced in contrast to their own face-to-face experiments.

Another very interesting series of replications and extensions was conducted by Cardenas (2000) using field laboratories in rural Colombia. Cardenas initially invited over 200 villagers to participate in a series of common-pool resource experiments. Several of his experiments closely paralleled the ones described above. Others extended the questions that could be addressed. The villagers who Cardenas invited were actual users of local forests for the extraction of firewood, natural fibres, and log timber as well as water. One of the basic questions he wanted to pursue was whether villagers, who were heavily dependent on local forests for wood products, would behave in a manner broadly consistent with that of undergraduate students at an American university.

The answer to this first question was positive. He wrote his instructions in Spanish and in a manner that would be easily understood by villagers. Instead of tokens, an easy medium for undergraduates, he asked villagers to decide on how many months a year they would spend in the forest gathering wood products as contrasted to using their time otherwise. Each villager had a copy of a payoff table, the same for all participants. It showed that as the number of months that any individual would spend in the forest, that individual would gain more returns. The return to all of them, however, depended on all of them keeping their harvesting time to certain levels.

In the baseline, no-communication experiments, Cardenas found a pattern similar to what we had found. Villagers substantially overinvested in the resource. While groups exhibited considerable variation, villagers on average achieved 57.7 percent of their optimal return in the last three rounds of the baseline experiments (Cardenas, 2000, p. 316). Face-to-face communication enabled them to increase efficiency on average to 76.1 percent of optimal. Considerable variation among groups existed, which Cardenas was able to explain using information about the participants filled in after that experiment was completed. He found, for example, that when most members of the group were already familiar with common-pool resources such as the collective use of a mangrove forest, they used the communication rounds more effectively than when most members of the group were dependent primarily on their own assets. Cardenas also found that 'social distance and group inequality based on the economic wealth of the people in the group seemed to constrain the effectiveness of communication for this same sample of groups' (Cardenas, 2000, p. 317; see also Cardenas, 2003).

4.3. Face-to-face communication among heterogeneous subjects

Hackett et al. (1994) conducted a series of CPR experiments to explore whether communication could ameliorate the problems identified in field settings related to heterogeneity among appropriators (Hardin, 1982; Johnson and Libecap, 1982; Libecap and Wiggins, 1984; Bardhan and Dayton-Johnson, 2002; Wiggins and Libecap, 1987). The task of agreeing to and sustaining agreements for efficient CPR appropriation is more difficult for heterogeneous appropriators because of the distributional conflict associated with alternative sharing rules. In heterogeneous settings, all appropriators may be made better off by adopting a new rule, but some will benefit more than others, depending upon the sharing rule chosen. Consequently, appropriators may fail to cooperate on the adoption of a sharing rule because they cannot agree upon what would constitute a fair distribution of benefits produced by cooperating.

In order to address appropriator heterogeneity, the Hackett et al. experimental design allows for two levels of input endowments. One subset of appropriators has large endowments of tokens (24); the other appropriators have small token endowments (8). Group allocations to the CPR at the asymmetric Nash equilibrium are greater than optimal, but not all rents from the CPR are dissipated.

In order for communication to enhance joint payoffs to a heterogeneous set of subjects, subjects must agree on (1) the target level of group allocations to the CPR and (2) a rule for allocating the target input allocation across appropriators. The existence of heterogeneity in endowments and in historic allocation levels has no effect on the first problem, but is likely to elicit disagreement over second problems.

Subjects knew with certainty the total number of decision makers in the group, their own token endowment and that of the others, the total number of tokens in the group, the transformation function, and the number of decision rounds in the current treatment condition. Subjects participated in two (consecutive) 10-round sequences of the asymmetric game. In the first 10 rounds, subjects were not allowed to communicate, but face-to-face communication was allowed during the second set. Prior to each 10-round treatment sequence, four subjects were assigned the 'large' token endowment, while the other four subjects were assigned the 'small' endowment. Two different mechanisms were used for assigning the large and small endowment positions: random and auction (based on Güth, 1988, 1994). In the first method, large endowments were assigned randomly. A multiple-unit ascending price auction was used as the alternative mechanism for assigning endowment positions because of its demand-revealing characteristics. For both treatment conditions with no-communication, they observed a level of rent accrual relatively close to that predicted by the Nash equilibrium (48.9%).

The opportunity to communicate led to a noticeable change in the pattern of allocations to the CPR. With the allocation rules agreed upon in communication rounds, subjects concentrated allocations to the CPR near the optimal allocation of 56 tokens in total. In the random-assignment and communication condition, individual CPR allocations of eight tokens represented the modal response (67%). In the auction and communication condition, however, Hackett et al. observed a spread of allocations clustered between 6 and 10 tokens. Thus, even in an environment of extreme heterogeneity in subject endowments, communication was a powerful mechanism for promoting coordination, resulting in rents very close to those observed in the homogeneous decision setting discussed above.

4.4. Face-to-face communication among only a subset of subjects

Three follow-up experimental designs were conducted by Schmitt et al. (2000). They used the same baseline design as first described with three variations in regard to the communication component of the situation. In all of their protocols for the communication aspects of the experiment (rounds 11 through 25), six out of the eight players were invited to communicate with one another in one location. In their first protocol, two of the remaining ‘players’ were computerized decision makers whose decisions were each the result of a random draw of a number between 4 and 12. The other six who could communicate with one another were informed about the constraints on the random draw for ‘Players 7 and 8.’ In their second protocol, two of the players were real subjects who had been separated from the other six players. The real subjects did not face a constraint on their decisions. In the third protocol, the two separated players were constrained to invest between 4 and 12 tokens in the common-pool resource. As in the baseline experiments, subjects were informed about the aggregate investment of all players after each round.

In all three protocols, subjects substantially improved their overall efficiency in the communication rounds as contrasted to the non-communication rounds. Constraining who could communicate to six out of the eight players did, however, make a difference when compared to designs where all subjects could communicate. The six communicating subjects were never certain whether a higher aggregated investment level than agreed upon reflected higher investments by the two ‘non-communicating’ players or whether some of the communicating group did not follow their own agreement. This uncertainty affected the capacity of the communicating group to make agreements in the first place. They could always blame the outsiders for any major overinvestment.

Major differences also occurred in the behavior across the three protocols. The ‘outsiders’ in Protocol 2 were the least constrained in their decisions. The members of the communicating group had the most difficulty in reaching agreements and following them in this design. The six communicating subjects in Protocol 2 had a ‘scapegoat’ they could blame for high investment levels in rounds 11 through 25. The subjects in Protocol 2 were less likely to come to an agreement in the first place, had a much higher deviation rate and size of deviation when they did agree, and obtained lower payoffs than in the earlier Ostrom et al. (1994) experiments or in Protocol 1 or 3. The problem of imperfect monitoring was less severe in Protocols 1 and 3. Some subjects among the communicating group were able to deviate without raising suspicion of cheating in these protocols, but they made only small increases over what the group had promised each other to do.

What this series of experiments found has considerable implications for those trying to achieve an agreement in the field to limit harvesting from a common-pool resource. The results provide evidence that ‘communication is less likely to be effective in preventing overharvesting in common-pool resource (CPR) environments in which a subset of appropriators either cannot or will not participate in collective action’ (Schmitt et al.). The lack of commitment by an outside group is not only a source of additional investment but also gives ‘insiders’ a scapegoat to blame if their own harvests are higher than agreed upon. ‘The problem becomes more severe when outsiders have less constraints on their overall appropriation behavior and their ability to behave strategically' (2000, p. 852).

5. Sanctioning experiments

In addition to communication among harvesters in most successful field settings, participants have also devised formal or informal ways of monitoring and sanctioning one another if rules are broken. In fact, the amazing speed with which irrigators responded to a perceived infraction while Ostrom was studying an irrigation system in Nepal led her to urge Roy Gardner and James Walker to help to develop a non-cooperative game-theoretical model and experimental design to study sanctioning in the lab. Engaging in costly monitoring and sanctioning behavior, even though often proposed as a way of providing selective incentives, was not consistent with the theory of norm-free, complete rationality (Elster, 1989, pp. 40–41). Thus, it was important to ascertain whether subjects in a controlled setting would actually make a costly decision in order to assess a financial punishment on the behavior of other participants as observed in the field. The short answer to this question is yes.

5.1. Initial sanctioning experiments

All sanctioning experiments used the 25-token design since appropriations had been much higher in this design. Subjects played ten rounds of the baseline game modified so that the individual contributions in each round were reported as well as the total outcomes. Subjects were then told that in the subsequent rounds they would have an opportunity to pay a fee in order to impose a fine on the payoffs received by another player. The fees ranged in diverse experiments from $.05 to .20 and the fines from $.10 to .80 (Ostrom et al., 1992). In brief, this series of experiments found that sanctioning occurs at far higher levels than the predicted level of zero.

Subjects react both to the cost of sanctioning and to the fee-to-fine relationships. Subjects assign more sanctions when the cost of sanctioning is lower and when the ratio of the fine to the fee is higher. Sanctioning is primarily directed at those who invested more in the CPR. In this set of experiments, subjects were able to increase their returns modestly from the CPR to 39 percent of maximum. When the costs of fees and fines were subtracted from the total, however, these gains were wiped out. On the other hand, when subjects were given a single opportunity to communicate prior to the implementation of sanctioning capabilities, they gained an average of 85 percent of the maximum payoffs (69 percent when the costs of the fees and fines were subtracted) (see Ostrom et al., 1992, 1994).

5.2. Replications of sanctioning experiments

Sanctioning experiments have also been replicated and extended by several scholars (Falk et al., 2002; Fehr and Gächter, 2000a,b) in other types of collective-action experiments. In a closely related experiment, Casari and Plott (2003) explored whether an institution that had been used in the Italian Alps for centuries and thought to be highly effective would generate positive results in a laboratory setting. The Alpine system had a relatively simple structure:

> The population of a village developed a contract among themselves, subject to the approval of the regional government, called 'Carte di Regola', where they described a system for monitoring and sanctioning those who are discovered violating or exceeding patterns of use

> that the villagers agreed upon in the contract. The 'Carte di Regola' specified in advance the conditions under which a sanction could be inflicted on a person found in violation of the contract and the amount of the fine... Any villager could report a violation but he usually incurred a cost in the form of a monitoring effort to discover the violator and additional costs to bring him to court. A share of such a fine usually went to the person who discovered the violator in order to give an incentive to monitor. (ibid, p. 218)

Casari and Plott used the same functional form for a payoff function as we had earlier used (see Eq. (1)), but they increased the monetary incentives by more than three-fold.[2] They first ran a baseline experiment that closely paralleled our earlier baseline experiments. Without communication or sanctioning, they replicated our earlier finding that the resource was substantially overused, even more than the Nash equilibrium. Subjects earned only 28.4 percent of the optimal return, while the Nash equilibrium would have earned them 39.5 percent. They also found substantial variations among individual subjects in the amount they overused the resource, as we had earlier.

Casari and Plott then changed the transformation function and the payoffs of the game. They used two sanctioning conditions, weak and strong. In both conditions, after the decision regarding harvesting has been made and the total investment levels have been made public, a subject could select an option to pay to inspect the decision made by any of the other subjects. After this decision, the harvesting decision of the inspected subject is made public information, but not the identity or number of subjects requesting an inspection. A fine is imposed for each unit appropriated above the announced level and transferred to the inspector.

In the experiments conducted with weak sanctions, slightly over half of the actions were inspected (a higher level than predicted by non-cooperative game theory). Subjects obtained closer to optimal levels of returns than they had without sanctions. In the strong sanction condition, the efficiency of the joint return was 94 percent, but when inspection fees are subtracted, the net return was 76.9 percent of optimal (2003, p. 238). Almost all actions were inspected, and it turned out that the lowest users were more aggressive inspectors than the highest users.

Casari and Plott found that the subjects behaved in a manner consistent with having heterogeneous preferences rather than all having the same preferences monotonically aligned with the payoff available. Some individuals appear to be more spiteful than others. This helps to explain the success of the 'Carte di Regola' system. The long-lasting field regime was able to use the 'heterogeneity of preferences to socially advantageous ends' (2003, p. 241). It channeled the preference of the more spiteful participants into socially useful purposes. Overall, they found that the experiment replicating the set of rules used in the Italian Alps greatly improved the efficiency of resource use as contrasted with the baseline experiment without the sanctioning options.

6. Endogenous rule making

In self-organized field settings, participants rarely impose exogenously authorized sanctions on one another and as we and Casari and Plott did in the sanctioning experiments. Sanctioning systems are more likely to emerge from an endogenous process of crafting their own rules, including the punishments that should be imposed if these rules are broken. Spending time and effort designing rules creates a public good for all of those involved. Crafting rules for an operational situation is thus a second-level dilemma that theorists have argued is no more likely to be solved than

[2] Casari and Plott also changed the instructions given to players so as to make the differences between the three conditions very distinct.

the original commons dilemma. The prediction that participants will not undertake such efforts is the reason given for repeated recommendation to impose rules and enforce them by external authorities. Since self-designed rules are found in many common-pool resource situations, it appears that participants frequently do design their own rules contrary to the theoretical prediction. Few scholars are able to witness these processes, however, in the field.

In order to observe what happens in these settings, subjects who had already experienced baseline and sanctioning experiments were recalled and given an opportunity to choose whether or not they would like to adopt a sanctioning mechanism, how much the fines and fees should be, and the joint investment strategy that they would like to adopt. All of these groups were endowed with 25 tokens in each round. Four out of six experimental groups adopted a covenant in which they specified the number of tokens they would invest and the level of fines to be imposed. The fines determined by the participants ranged in size from $.10 to 1.00 (Ostrom et al., 1992).

The groups that crafted their own agreements were able to achieve an average of 93 percent of the maximum in the periods after their agreement, and the defection rate for these experiments was only 4 percent. The two groups that lacked agreement did not fare nearly as well. They averaged 56 percent of the maximum available returns and faced a defection rate of 42 percent (Ostrom et al., 1992). Consequently, those subjects who used an opportunity to covenant with one another to agree on a joint strategy chose their own level of fines and received very close to optimal results based entirely on their own promises and their own willingness to monitor and sanction one another when it was occasionally necessary. Other experiments have examined endogenous voting patterns under different rules (Walker et al., 2000).

7. Exogenous rule making

Cardenas et al. (2000) report on five experiments conducted in Columbia where subjects were told that a new rule would go into force that mandated that they should spend no more than the *optimal* level of time in the forest each round. They were told that there would be a 50 percent chance that conformance to the rule would be monitored each round. The experimenter rolled a die in front of the subjects each round to determine whether an inspection would take place. If an even number showed up, there would be an inspection. The experimenter then drew a number from chits (numbered between 1 and 8) placed in a hat to determine who would be inspected. Thus, the probability that anyone would be inspected was 1/16 per round—a realistic probability for monitoring forest harvesting in rural areas. The monitor checked the investment of the chosen subject. If the person was over the limit imposed, a penalty was subtracted from the payoff to that person. No statement was made to others whether the appropriator was complying with regulations or not.

The subjects in this experimental condition actually *increased* their harvesting levels in contrast to experiments when the subjects could simply communicate with each other on a face-to-face basis (see also Beckenkamp and Wojtyniak, 1998). What was remarkable about these experiments was that subjects who communicated with one another were able to achieve a higher joint return than the subjects who had an optimal but imperfectly enforced external rule imposed on them. As the authors conclude

> We have presented evidence that indicates that local environmental policies that are modestly enforced, but nevertheless are predicted by standard theory to be welfare-improving, may be ineffective. In fact, such a policy can do more harm than good, especially in comparison to allowing individuals collectively to confront local environmental dilemmas without

> intervention. We have also. . .presented evidence that the fundamental reason for the poor performance of external control is that it crowded out group-regarding behavior in favor of greater self-interest. (Cardenas et al., 2000, p. 1731)

8. Conclusion

In light of this short overview, let me return to the question regarding the value-added of experimental research to the overall research program on the impact of diverse institutions on behavior and outcomes of users of common-pool resources. First, we and others have shown that solving these problems is non-trivial, but feasible for those directly involved if they can communicate with one another and agree on future actions. Without communication of some sort, however, subjects in the laboratory behave in the aggregate similar to the predicted Nash equilibrium of the non-cooperative game. When they cannot affect each other's expectations and trust, individuals behave as short-term, payoff maximizers. When asked whether easy solutions are likely to be achieved for problems involving large, amorphous groups that face significant problems of communicating, such as the overuse of ocean fisheries or global warming, I always respond in the negative (Ostrom et al., 1999). These are far more challenging problems than those faced by smaller groups of users who frequently assert the authority to make their own rules.

Second, we have shown that groups who can communicate and design their own rules achieve substantial improvements. Third, subjects who have rules imposed on them (even ones that specify an *optimal* appropriation level for their environment) begin to cheat on one another relatively rapidly. With external monitoring that approximates what is found in many rural areas, such groups achieve far less than their own 'cheap talk' agreements. Fourth, the experiments have shown that face-to-face communication can also help participants overcome problems of heterogeneity and subgroup divisions.

Thus, at a more general level, our experiments, along with field research and theoretical efforts, lead us to posit that the crucial variables to enhance cooperation in regard to common-pool resources and other forms of collective action are those that enhance reciprocity, individual reputations, and trust (Ostrom, 1998). Repeated interactions where participants gain trust that others are trustworthy and engage in reciprocal relationships lead to high levels of performance. Field research shows that many rules adopted by local resource users (and sometimes denigrated by policy analysis) affect reciprocity, reputations, and trust. Large groups with the authority to make their own rules tend to create nested decision-making units so that smaller units can engage in effective communication and decision making about aspects of a smaller subsystem (Ostrom, 1990). There are, however, no single rules guaranteed to have a positive effect in all settings (Ostrom, 2005).

Unfortunately, some policy advisors have thought that involving the users of a resource in some kind of participatory activity is an easy way to overcome resistance to external programs designed to protect resources. This is *not* the lesson we have learned. Calling resource users to a single meeting and asking them 'to participate' while telling them what a project will do, is just an exogenous change that is likely to crowd out positive endogenous processes (Frey, 1994). These efforts are unlikely to create a setting in which reciprocity and trust can be achieved.

Acknowledgments

An earlier version of this article was presented at the conference on 'Experiments in Natural Resource Economics: Fisheries and Energy,' Akureyri, Iceland, 20–23 May 2004. Sections of the article draw on Chapter 3 of Ostrom (2005). The author is appreciative of funding from the National Science Foundation, the useful comments and suggestions by two anonymous reviewers, and the outstanding editing by David Price and Patty Lezotte.

References

Acheson, J.M., 2003. Capturing the Commons: Devising Institutions to Manage the Maine Lobster Industry. University Press of New England, Hanover, NH.

Acheson, J.M., Gardner, R., 2004. Strategies, conflict, and the emergence of territoriality: the case of the Maine lobster industry. American Anthropologist 6, 296–307.

Agrawal, A., 1999. Greener Pastures: Politics, Markets and Community among a Migrant Pastoral People. Duke University Press, Durham, NC.

Bardhan, P., Dayton-Johnson, J., 2002. Unequal irrigators: heterogeneity and commons management in large-scale multivariate research. In: Ostrom, E., Dietz, T., Dolšak, N., Stern, P., Stonich, S., Weber, E., National Research Council, Committee on the Human Dimensions of Global Change (Eds.), The Drama of the Commons. National Academy Press, Washington, DC, pp. 87–112.

Beckenkamp, M., Wojtyniak, B., 1998. Optimizing fines in commons dilemmas. Discussion Paper, University of Karlsruhe.

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.

Cardenas, J.-C., 2003. Real wealth and experimental cooperation: evidence from field experiments. Journal of Development Economics 70, 263–289.

Cardenas, J.-C., Stranlund, J.K., Willis, C.E., 2000. Local environmental control and institutional crowding-out. World Development 28, 1719–1733.

Casari, M., Plott, C.R., 2003. Decentralized management of common property resources: experiments with a centuries-old institution. Journal of Economic Behavior and Organization 51, 217–247.

Deadman, P., 1997. Modeling individual behavior in common pool resource management experiments with autonomous agents. Ph.D. Dissertation, University of Arizona, Tucson.

Dietz, T., Ostrom, E., Stern, P., 2003. The struggle to govern the commons. Science 302, 1907–1912.

Elster, J., 1989. The Cement of Society: A Study of Social Order. Cambridge University Press.

Falk, A., Fehr, E., Fischbacher, U., 2002. Appropriating the commons: A theoretical explanation. In: Ostrom, E., Dietz, T., Dolšak, N., Stern, P., Stonich, S., Weber, E., National Research Council, Committee on the Human Dimensions of Global Change (Eds.), The Drama of the Commons. National Academy Press, Washington, DC, pp. 157–191.

Fehr, E., Gächter, S., 2000a. Cooperation and punishment in public goods experiments. American Economic Review 90, 980–994.

Fehr, E., Gächter, S., 2000b. Fairness and retaliation: the economics of reciprocity. Journal of Economic Perspectives 14, 159–181.

Frey, B.S., 1994. How intrinsic motivation is crowded out and in. Rationality and Society 6, 334–352.

Gibson, C., McKean, M., Ostrom, E. (Eds.), 2000. People and Forests: Communities, Institutions, and Governance. MIT Press, Cambridge, MA.

Gibson, C., Williams, J.T., Ostrom, E., 2005. Local enforcement and better forests. World Development 33, 273–284.

Gordon, H.S., 1954. The economic theory of a common property resource: the fishery. Journal of Political Economy 62, 124–142.

Güth, W., 1988. On the behavioral approach to distributive justice—a theoretical and experimental investigation. In: Maital, S. (Ed.), Applied Behavioral Economics, vol. 2. New York University Press, New York, pp. 703–717.

Güth, W., 1994. Distributive justice: a behavioral theory and empirical evidence. In: Brandstatter, H., Guth, W. (Eds.), Essays on Economic Psychology. Springer-Verlag, Berlin, pp. 153–176.

Hackett, S., Schlager, E., Walker, J., 1994. The role of communication in resolving commons dilemmas: experimental evidence with heterogeneous appropriators. Journal of Environmental Economics and Management 27, 99–126.

Hardin, G., 1978. Political requirements for preserving our common heritage. In: Brokaw, H.P. (Ed.), Wildlife and America. Council on Environmental Quality, Washington, DC, pp. 310–317.

Hardin, R., 1982. Collective Action. Johns Hopkins University Press, Baltimore, MD.

Jager, W., Janssen, M., 2002. Using artificial agents to understand laboratory experiments of common pool resources with real agents. In: Janssen, M. (Ed.), Complexity and Ecosystem Management. Edward Elgar, Cheltenham, UK, pp. 75–102.

Jager, W., Janssen, M., Vlek, C.A.J., 2002. How uncertainty stimulates over-harvesting in a resource dilemma: three process explanations. Journal of Environmental Psychology 22, 247–263.

Johnson, R.N., Libecap, G.D., 1982. Contracting problems and regulation: the case of the fishery. American Economic Review 72, 1005–1023.

Kaul, M.C., 1996. Common Lands and Customary Law: Institutional Change in North India over the Past Two Centuries. Oxford University Press, New Delhi.

Lam, W.F., 1998. Governing Irrigation Systems in Nepal: Institutions, Infrastructure and Collective Action. ICS Press, Oakland, CA.

Libecap, G.D., Wiggins, S.N., 1984. Contractual responses to the common pool: prorationing of crude oil production. American Economic Review 74, 87–98.

Moran, E., Ostrom, E., 2005. Seeing the Forests and the Trees: Human–Environment Interactions in Forest Ecosystems. MIT Press, Cambridge, MA.

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

Ostrom, E., 1992. Crafting Institutions for Self-governing Irrigation Systems. ICS Press, San Francisco, CA.

Ostrom, E., 1998. A behavioral approach to the rational choice theory of collective action. American Political Science Review 92, 1–22.

Ostrom, E., 2005. Understanding Institutional Diversity. Princeton University Press, Princeton, NJ.

Ostrom, E., Burger, J., Field, C., Norgaard, R., Policansky, D., 1999. Revisiting the commons: local lessons, global challenges. Science 284, 278–282.

Ostrom, E., Gardner, R., Walker, J., 1994. Rules, Games and Common-pool Resources. University of Michigan Press, Ann Arbor.

Ostrom, E., Walker, J., 1991. Communication in a commons: cooperation without external enforcement. In: Palfrey, T.R. (Ed.), Laboratory Research in Political Economy. University of Michigan Press, Ann Arbor, pp. 287–322.

Ostrom, E., Walker, J., Gardner, R., 1992. Covenants with and without a sword: self-governance is possible. American Political Science Review 86, 404–417.

Poteete, A., Ostrom, E., 2004. Heterogeneity, group size and collective action: the role of institutions in forest management. Development and Change 35, 435–461.

Rocco, E., Warglien, M., 1995. Computer Mediated Communication and the Emergence of Electronic opportunism. Department of Economics, Laboratory of Experimental Economics, University of Venice, Italy.

Sally, D., 1995. Conversation and cooperation in social dilemmas: a meta-analysis of experiments from 1958 to 1992. Rationality and Society 7, 58–92.

Schlager, E., 1994. Fishers' institutional responses to common-pool resource dilemmas. In: Ostrom, E., Gardner, R., Walker, J. (Eds.), Rules, Games, and Common-pool Resources. University of Michigan Press, Ann Arbor, pp. 247–266.

Schmitt, P., Swope, K., Walker, J., 2000. Collective action with incomplete commitment: experimental evidence. Southern Economic Journal 66, 829–854.

Scott, A.D., 1955. The fishery: the objectives of sole ownership. Journal of Political Economy 63, 116–124.

Shivakoti, G.P., Ostrom, E. (Eds.), 2002. Improving Irrigation Governance and Management in Nepal. ICS Press, Oakland, CA.

Simon, A., Gorgura, H., 2003. Say the magic word: effective communication in social dilemmas. Working Paper. University of Washington, Department of Political Science, Seattle.

Tang, S.Y., 1992. Institutions and Collective Action: Self-Governance in Irrigation. ICS Press, San Francisco, CA.

Varughese, G., Ostrom, E., 2001. The contested role of heterogeneity in collective action: some evidence from community forestry in Nepal. World Development 29, 747–765.

Walker, J.M., Gardner, R., Herr, A., Ostrom, E., 2000. Collective choice in the commons: experimental results on proposed allocation rules and votes. The Economic Journal 110, 212–234.

Wiggins, S.N., Libecap, G.D., 1987. Firm heterogeneities and cartelization efforts in domestic crude oil. Journal of Law, Economics, and Organization 3, 1–25.

Yandle, T., Dewees, C., 2003. Privatizing the commons... 12 years later: fishers' experiences with New Zealand's market-based fisheries management. In: Dolšak, N., Ostrom, E. (Eds.), The Commons in the New Millennium: Challenges and Adaptations. MIT Press, Cambridge, MA, pp. 101–127.