Journal of Personality and Social Psychology
1986, Vol. 51, No. 1, 110-116

0022-3514/86/$00.75

The Provision of a Sanctioning System as a Public Good

Toshio Yamagishi
University of Washington

Both the rational-structural approach and the goal/expectation approach to the problem of public goods have theoretical difficulties. The structural approach requires the provision of a sanctioning system to solve the free rider problem. However, a sanctioning system is also a public good because its benefits can be enjoyed by all members regardless of their contribution to its provision. A new problem of the same kind is thereby created in the process of solving the original public good problem. The goal/expectation approach assumes the inducement of other members to mutual cooperation through individuals' cooperative actions, a situation which will be almost impossible in larger groups. To overcome these theoretical difficulties in the existing approaches, a new approach called the structural goal/expectation approach is proposed. According to this new approach, members who have realized the undesirable consequence of free riding and the importance of mutual cooperation will cooperate to establish a sanctioning system which assures other members' cooperation instead of trying to induce other members into mutual cooperation directly through cooperative actions. One important condition for their voluntary cooperation in the establishment of a sanctioning system is their realization that voluntarily based cooperation is impossible. Predictions derived from the new approach are supported in an experiment using 48 four-person groups.

Whenever members of a group or a society have free access to a common property, whether it is public broadcasting, national security, crime control, or a clean living room in a shared house, free riding is a potential problem. When people do not need to contribute to the provision and maintenance of a common property to get access to it, it is rational for each individual to use it without paying the cost for its provision and maintenance. However, if everybody behaves in this rational fashion, nobody will bear the cost for its provision/maintenance and thus everybody will lose the opportunity to enjoy the benefit associated with its use. This problem of public goods (or the problem of social dilemmas) and free riding has recently received increasing attention in the social sciences, including sociology, economics, political science, and psychology. Despite the diversity of these fields and the many terms used to conceptualize the problem—social traps, commons dilemmas, *n*-person prisoner's dilemmas, public goods, free riding, externality, social dilemmas, and so forth—theoretical approaches to this problem may be largely classified into the following: one emphasizing the rational nature of individuals and the other emphasizing the more psychological aspects of the problem. The purpose of the present study is (a) to clarify the logical limitations of the current approaches, (b) to present a new approach intended to overcome these logical limitations, and (c) to test some of the predictions derived from the new approach in an experiment using small groups.

According to the first approach (cf. Olson, 1965), an individual actor is assumed to contribute to the provision/maintenance of a public good only when the extra benefit associated with his or her contribution exceeds the cost. Insofar as we assume that people are rational in this fashion, the only solution to the public goods problem is in the modification of the incentive (or payoff) structure, so that the personal benefit from contribution exceeds the personal cost. A large number of studies of public goods and social dilemmas have investigated various forms of structural changes and have repeatedly demonstrated the effectiveness of this type of solution. The structural factors manipulated in these studies include cost of contribution (or the temptation for defection) and/or the size of the public good (or the gain of cooperation; Bonacich, Shure, Kahan, & Meeker, 1976; Goehring & Kahan, 1976; Kelley & Grzelak, 1972; Komorita, Sweeney, & Kravitz, 1980; Marwell & Ames, 1979, 1980), introduction of positive or negative sanctions (or selective incentives; Caldwell, 1976; Linaweaver, Geyer, & Wolff, 1967; Maki, Hoffman, & Berk, 1978; Sato, 1984, 1985), territorialization or privatization (Acheson, 1975; Cass & Edney, 1978; Messick & McClelland, 1983), group size (Bonacich et al., 1976; Fox & Guyer, 1977; Hamburger, Guyer, & Fox, 1975; Komorita & Lapworth, 1982; Marwell & Schmitt, 1972), establishment of leadership (Messick et al., 1983; Samuelson, Messick, Rute, & Wilke, 1984), and so on.

Despite empirical evidence for the effectiveness of a structural solution to the public goods problem, it has serious problems. First, some structural changes are very difficult or even impossible to implement because of physical or technical limitations and/or political and other problems. Furthermore, and more importantly, even if the needed structural changes are technically feasible and no political opposition is expected,

This research was supported by Japanese Ministry of Education Scientific Research Grant 5810040. I wish to thank Yasuhisa Hama, Hayuru Ito, and Hiroko Takeda for their assistance, and Mary Brinton, David Messick and Michael Hechter for their comments on the earlier drafts of this article.

Correspondence concerning this article should be addressed to Toshio Yamagishi, Department of Sociology, University of Washington, Seattle, Washington 98195.

there remains the problem of who shall bear the cost required for the change. Because the needed structural changes benefit all members of a group regardless of whether they paid the costs for the change or not, the implementation of the needed change involves another public good, which may be called the *second-order public good* (cf. Oliver, 1980). Not contributing to the implementation of the change is then the rational choice, and as a result the second-order public good (or the implementation of the needed change) will fail to be provided. In brief, insofar as people are assumed to be rational, structural changes are needed to solve a public goods problem; but the needed structural changes may produce the second-order public goods problem regarding the implementation of the needed structural change.

The second approach (cf. Brubaker, 1975; Pruitt & Kimmel, 1977) assumes that people have a broader perspective than assumed in the rational-structural approach. In many public goods situations, people in fact know about the disastrous consequences of their individually rational actions. And, people often know that cooperation (especially the cooperation of others) is essential to their long-term benefit and that individuals' uncooperative actions lead to disastrous consequences. However, even if they realize the importance of cooperation many people still do not cooperate. Why? The goal/expectation theory proposed by Pruitt and Kimmel (1977) summarizes the answer to this question. According to Pruitt and Kimmel (1977), cooperation in experimental dilemma games often results from long-range thinking that is directed "to achieve the goal of establishing and/or maintaining continued mutual cooperation" (p. 375). However, the development of the goal of attaining continued mutual cooperation is not a sufficient condition for cooperation. Given the establishment of this goal, the decision of cooperation/defection still depends on the trust that partners will reciprocate cooperation; members having developed the goal of mutual cooperation will actually cooperate only when they are convinced that partners will reciprocate their initiative and will not exploit their goodwill. Unless members have this trust, the risk is too strong that their cooperation may be exploited by the partners.

The effect that trust in other members has on enhancing cooperation in social dilemmas has been demonstrated in many studies (Alcock & Mansell, 1977; Fox & Guyer, 1977; Messick et al., 1983; Sato & Yamagishi, in press). However, these studies used only groups of relatively small size, and whether or not trust improves cooperation in large groups is uncertain. There is at least one important factor which could limit the applicability of goal/expectation theory to small groups. According to the theory, sustained mutual cooperation is achieved when each actor is convinced that his or her initiative in cooperation will be reciprocated. With this conviction of reciprocity (or trust in the partners), each actor voluntarily cooperates, and as a result each actor's conviction or trust is substantiated. This process may work in a dyadic situation or in small groups where it is relatively easy for members to feel interdependency among partners. However, in large groups, a single member's action has very low visibility or impact on other members. That is, an individual's action is hardly noticed by other members, and it has practically no effect on other members' decisions/actions. It is almost impossible to direct others' decisions, through one's own actions, into mutual cooperation. Even those who have developed the goal of mutual cooperation will hesitate to cooperate because unilateral cooperation involves a large risk of being exploited. This will make the application of goal/expectation theory less appropriate in large groups. In fact, Fox and Guyer's (1977) study shows that expectations of other members' behavior (whether they are cooperative or not) has a smaller effect in larger groups.

The logical difficulties of the two approaches discussed earlier can be avoided, and one logically consistent and widely applicable solution can be provided when it is assumed that people who have developed the goal of mutual cooperation do cooperate for the implementation of the needed structural change (let us term this *instrumental cooperation*) rather than simply engaging in cooperative actions in the original public good situation (let us term this *elementary cooperation*). This new approach, which is termed the *structural goal/expectation* (SG/E) approach, shares the first assumption or requirement for cooperation with the original goal/expectation theory. That is, the necessary condition for voluntarily based structural change (or instrumental cooperation) is developing the goal of attaining and maintaining mutual cooperation. Given the development of the goal of mutual cooperation, however, the sufficient condition for the actual cooperative behavior takes different forms in the SG/E approach and in the original goal/expectation theory. In the latter, *mutual trust* is the key to actual cooperation. However, prevailing mutual trust that other members will voluntarily cooperate may prevent people from realizing the necessity of the structural change. If people believe that others are also cooperative, they may not realize the importance of changing the structure itself. For example, insofar as people believe that others can handle handguns responsibly and will not abuse them, the necessity of legislative actions for controlling handguns will not be realized. Or, if people believe that social welfare can be achieved simply through voluntary donations, government actions will not be called for. Thus, in the SG/E approach, the sufficient condition for actual instrumental cooperation (for the implementation of the needed change) is that people realize the futility of voluntarily based cooperation and the effectiveness and importance of structural changes.

By adopting the basic assumption in the goal/expectation approach that people are willing to cooperate when they realize its importance, the new approach can avoid the difficulty in the original structural approach. That is, the vicious cycle in the original structural approach is removed because people are assumed to voluntarily cooperate in the implementation of the needed structural changes. At the same time, this approach can be applied to large groups because the structural changes, when they are implemented, will function to improve other members' cooperation. In sum, this approach synthesizes the structural approach and the goal/expectation approach, and emphasizes not only the direct effect of voluntarily based structural changes on cooperation, but also their indirect effects through assuring other members' cooperation.

The Experiment

Several predictions derived from the SG/E approach are tested in this experiment. In this experiment, each member of

a 4-person group was given resource money which they could keep to themselves or contribute to the provision of a public good. In order to retain important properties of larger groups, the increase in the personal benefit due to own contribution was reduced to zero, and subjects were not given chances to see each other in person (see the Procedure section for details). In addition, some groups were given opportunities to develop a negative sanctioning system which punishes the least cooperative member in the group. The level of punishment depended on the total amount of contribution made by the group members to the sanctioning system (separate from the contribution to the original public good). Thus, the sanctioning system itself was also a public good. Besides these experimental manipulations, two types of groups were used in the experiment, consisting of subjects who scored high and low, respectively, in regard to trusting other people.

Method

Predictions

Prediction 1. When an opportunity to develop a sanctioning system does not exist, the SG/E approach as well as the original goal/expectation approach predicts that elementary cooperation will be greater among high-trusters (HTs) than among low-trusters (LTs), because the group size is small. The structural approach provides no predictions concerning the effect of trust.

Prediction 2. When an opportunity to develop a sanctioning system exists, LTs will be more likely to realize the importance of developing such a system than HTs. Then, LTs will contribute to the sanctioning system more than HTs. This prediction is unique to the SG/E approach. Because the original goal/expectation theory does not distingiush the two types of cooperation (elementary and instrumental), it predicts the opposite pattern to the one predicted by the SG/E approach: that is, more instrumental cooperation by HTs than by LTs.

Prediction 3. The SG/E approach as well as the structural approach predicts more elementary cooperation when a sanctioning system exists. The original goal/expectation theory provides no predictions concerning the effect of the sanctioning system.

Prediction 4. In addition to this main effect of the sanctioning system (Prediction 3), the SG/E approach predicts that the effect of a sanctioning system on elementary cooperation will be stronger among LTs than among HTs. This is because LTs will be more instrumentally cooperative than HTs (Prediction 2), and as a result the sanctioning system will have a stronger impact in the LT condition than in the HT condition.

In sum, the greater instrumental cooperation by LTs (Prediction 2) and the interaction between sanctioning system and trust (Prediction 4) are the predictions unique to the SG/E approach. Neither the structural approach nor the goal/expectation approach can provide these two predictions. It should be noted, however, that support for these predictions does not imply that either of the other two approaches are wrong. Rather, it would imply that the other two approaches are not sufficient by themselves; they need to be synthesized to produce the predictions.

Subjects

A total of 192 undergraduate students in 48 same-sex 4-person groups (24 male and 24 female groups) were used as subjects. First, a pool of students of about 600 were recruited from various classes at Hokkaido University and other colleges in Sapporo, Japan. At the time of recruitment, these students filled out a five-item questionnaire designed to measure their general level of trust, as well as other information needed to schedule them to the experimental session. Monetary rewards were emphasized at the time of recruitment. The pool was then divided into three groups according to their trust level: the low-trust group (the lowest 1/3 on the trust score), the medium-trust group (the medium 1/3), and the high-trust group (the highest 1/3). Actual participants were selected from the LT group and the HT group (96 LT and 96 HT students). In selecting participants from these two groups, it was verified that participants of the same experimental session did not know each other. The actual experimental session took place about 3 to 6 weeks after the recruitment. Thus, the pretest measurement of trust was considered to have only minor effects on subjects' decisions at the time of the experiment, if it had any effect at all.

Preexperimental Questionnaire

The questionnaire designed to measure subjects' general level of trust consisted of the following five items: (1) Most people tell a lie when they can benefit by doing so. (2) Those devoted to unselfish causes are often exploited by others. (3) Some people do not cooperate because they pursue only their own short-term self-interest. Thus, things that can be done well if people cooperate often fail because of these people. (4) Most people are basically honest (reversed item). (5) There will be more people who will not work if the social security system is developed further. Items 3 and 5 were taken from the "fear" scale used in Yamagishi and Sato (1986), and the other items were taken from the trust scale used in Sato and Yamagishi (in press). This 5-item trust scale was originally developed by Yukawa (1985) based on the factor analysis of 60 items related to trust. The current trust scale represents two most important factors observed in the factor analysis, one of which concerns the belief that people are basically honest (Items 1 and 4), and the other of which concerns the belief that trusting others is risky (Items 2, 3, and 5). Past research (Sato & Yamagishi, in press; Yamagishi & Sato, 1986) has demonstrated that the items used in the scale successfully predict subjects' cooperation level when no sanctioning exists.

Design

The design consisted of three between-subjects factors: (a) sanctioning (no sanctioning, low sanctioning, and high sanctioning), (b) trust (low trust and high trust), and (c) sex. In addition, the design included trial sessions as a repeated factor. The behavior of the subjects within each group was interdependent, so that the test of these factors requires groups rather than individual members as the unit of analysis.

Procedure

The experiment was run in groups of 4 subjects. As soon as each subject arrived at the laboratory, he or she was immediately led to a separate compartment so that he or she had no personal contact with other subjects. Each compartment was equipped with a microcomputer (Sharp MZ700) connected to a host computer (AiM16). All relevant information was relayed to the subject via the compuer display, and the subject's decisions were relayed to the host computer and recorded there.

The experiment consisted of 12 sessions. In each session, subjects were first notified that they were each given 100 yen (about 40 cents), and then were asked how much of the 100 yen they wanted to contribute to other members of the group (in steps of 10 yen). They could keep the rest of the money to themselves. The amount contributed was then doubled in value and distributed equally among the other 3 members. Thus, each member's contribution produced no increase in his or her own benefit; only other members could benefit. This procedure was utilized, instead of the more usual procedure of allocating a public good equally to all members, in order to make the experimental situation comparable to larger group situations where each individual's contribu-

tion has practically no effect on his or her own benefit. Thus, it was rational for each individual not to contribute; but each could receive 200 yen, $(100 \times 3 \times 2)/3$, if all contributed all of their money in contrast to 100 yen if none contributed.

After deciding how much of the 100 yen they wanted to contribute for the bonus, subjects in either the *low-sanctioning* (LS) or the *high-sanctioning* (HS) condition were again asked to decide how much they wanted to contribute to the "punishment fund." They could contribute any amount between 0 yen and 100 yen (in steps of 10 yen), regardless of the amount of contribution for the bonus. They paid this amount from their cumulative earnings. (In case they had not accumulated enough money, they could borrow from the experimenter.) Subjects in the *no-sanctioning* (NS) condition were not asked to make this decision. In the LS or HS condition, the total amount of money donated to the punishment fund by the members (LS condition) or twice that amount (HS condition) was subtracted from the cumulative earnings of the subject whose contribution level to the bonus (not including the contribution to the punishment fund) was the lowest among the 4 members. Thus, if the 4 members, A, B, C, and D, in the HS condition contributed 80 yen, 50 yen, 30 yen, and 100 yen, respectively, to the provision of the bonus, and 10 yen, 0 yen, 30 yen, and 20 yen, respectively, to the punishment fund, A's earnings = 100 (the amount given by the experimenter) − 80 (the amount of contribution for the bonus) + $(50 \times 2)/3$ (bonus provided by B's contribution) + $(30 \times 2)/3$ (bonus provided by C's contribution) + $(100 \times 2)/3$ (bonus provided by D's contribution) = 140 yen. Because C's contribution was the lowest among the 4 members, twice the total of all 4 members' contribution to the punishment fund, $2 \times (10 + 0 + 30 + 20) = 120$ yen, was subtracted from his or her earnings, so that his or her overall earnings in this session = $100 - 30 + (80 \times 2)/3 + (50 \times 2)/3 + (100 \times 2)/3 - 120 = 103$ yen. When there were two or more members whose contribution levels were the same and the lowest, these members suffered an equal share of the punishment (the total amount of punishment divided by the number of the lowest contributors). When all 4 members contributed 100 yen, none was punished.

Once all 4 members decided how much they wanted to contribute for the bonus (NS condition) and how much they wanted to contribute to the punishment fund as well (LS and HS conditions), each member received on his or her screen display (1) how much he or she contributed for the bonus and the punishment fund (in LS and HS conditions), (2) the total amount of contributions made by the other members and the total benefit from these contributions (the amounts earned by individual members were not told to the subjects), and (3) the total earnings in this session up to this step, that is, 100 yen − the amount in Step 1 + the amount in Step 2. Subjects in LS and HS conditions were then informed of (4) the total amount of contributions to the punishment fund, (5) how many members received punishment, (6) the amount of punishment each of these lowest contributors received, and (7) whether the subject himself or herself was the lowest contributor or not. When the subject was the lowest contributor (or one of the lowest contributors), he or she was informed that the amount of punishment just indicated was subtracted from the amount shown in Step 3. Finally, subjects in all conditions were informed of their own cumulative earnings up to this trial.

The instruction sheet contained information about the experiment described earlier and some examples of how earnings are determined. However, subjects were not told how many sessions there would be, to avoid the occurrence of behavior specific to the final few trials. After all members finished reading the instructions and all questions were answered, the experiment started. Each trial took less than 2 min on the average. After the 12th trial, subjects filled out a postexperimental questionnaire. Then, they were paid individually according to their earnings in the experiment and were dismissed individually. Subjects' earnings were generally between 1,000 yen (about $4) and 2,000 yen (about $8). The whole experiment including instructions and postexperimental questionnaire took about 1 hr.

Results

Because the results of a preliminary analysis revealed no main or interaction effects involving sex either on elementary or instrumental cooperation, sex was excluded from the following analysis as a factor. Twelve trial sessions were aggregated into three trial blocks of four trials each. (The use of 2, 4, and 6 blocks did not produce any detectable differences in the analysis.)

Predictions

Prediction 1. As shown in Figure 1, the HT's elementary cooperation level (55.65) was higher than the LT's (33.18) in the NS condition. The simple main effect of trust in this condition in the Trust × Sanctioning × Trial Block analysis of variance (ANOVA) of elementary cooperation was used to test this prediction, and the effect was significant, $F(1, 42) = 9.55$, $p < .01$. Thus, Prediction 1 was clearly supported.

Prediction 2. The average level of contribution to the punishment fund (instrumental cooperation) is also shown in Figure 1. LTs (13.70) contributed more to the punishment fund than

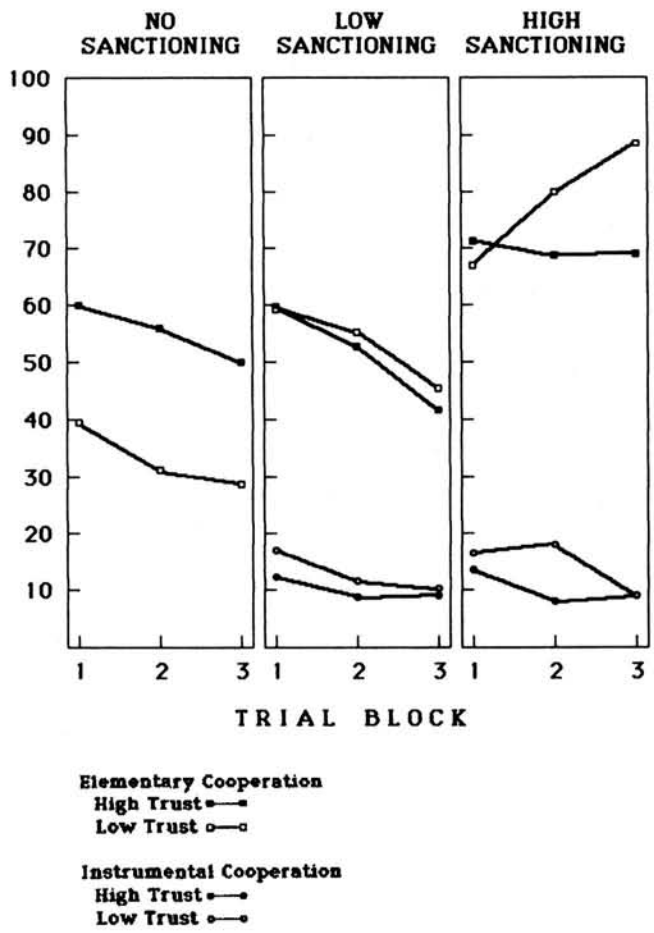

Figure 1. Average elementary and instrumental cooperations.

HTs (10.08), and the main effect of trust in the Trust × Sanctioning × Trial Block ANOVA was significant, $F(1, 28) = 5.30$, $p < .05$. Thus, the prediction based on the SG/E approach was clearly supported, and the prediction based on the original goal/expectation theory was rejected.

Prediction 3. The general level of elementary cooperation was the highest in the HS condition (74.10), followed by the LS condition (52.36), and finally by the NS condition (44.41). The main effect of sanctioning in the Trust × Sanctioning × Trial Block ANOVA was highly significant, $F(2, 42) = 17.89$, $p < .001$. Thus, Prediction 3 was clearly supported.

Prediction 4. As predicted, the Trust × Sanctioning interaction effect was significant, $F(2, 42) = 5.10$, $p < .01$. The simple main effect of sanctioning was significant both in the HT condition, $F(2, 42) = 3.48$, $p < .05$, and in the LT condition, $F(2, 42) = 19.52$, $p < .001$, but sanctioning explained 5.6 times as much sum of squares (SS) in the LT condition ($E^2 = 39.6/112.1 = .35$) than in the HT condition ($E^2 = 7.0/112.1 = 0.06$). Thus, Prediction 4 was also clearly supported.

Other Findings

In addition to supporting these predictions, the results of this experiment revealed several interesting patterns. First, trial block interacted with sanctioning, $F(4, 84) = 7.20$, $p < .001$, but not with trust in the ANOVA of elementary cooperation. Another significant effect in this ANOVA was the main effect of trial block, $F(2, 84) = 3.73$, $p < .05$. The other effects were not significant. Post hoc tests of the simple main effect of trial block revealed that the effect of trial block was significant in all Trust × Sanctioning Cells, except in the HS–HT cell. (Because these tests were all post hoc, and there are 37 possible simple main effects in this ANOVA, $[.05 \times 7]/37 = .01$ was used as the significance level in these tests of simple main effects.) These simple main effects indicate that subjects reduced their elementary cooperation over time when sanctioning did not exist or was weak, regardless of the level of their trust. However, when the sanctioning was strong, they increased their elementary cooperation (LT) or at least kept it at the relatively high level (HT).

These results seem to suggest that the mere existence of the sanctioning opportunity was enough to scare subjects about the possible punishment, so that they increased their elementary cooperation in the early trials. However, both the direct (punishment) and the indirect (improved expectations of other members' cooperation) effects of the sanctioning were not strong enough when the sanctioning was weak, so that members began defecting over time. On the other hand, when the sanctioning was strong, these effects of sanctioning were enough to inhibit subjects from defecting over time. The difference between the HT condition and the LT condition may be explained by the difference in the instrumental cooperation between the two conditions. That is, because LTs contributed more to the punishment fund than HTs, the effects of sanctioning were even stronger among LTs than HTs; the effects were barely strong enough in the HT condition to keep subjects from defecting more, whereas they were strong enough in the LT condition to make subjects contribute more.

Another interesting finding in this experiment is that the difference in the instrumental cooperation between HTs and LTs almost disappeared in the last trial block, especially in the HS condition (see Figure 1). Despite this drop in instrumental cooperation in the LT condition during the last trial block, LT's elementary cooperation increased substantially during the same trial block. This would imply that, once the efficacy of the sanctioning system is felt by the group members during the early trials, strong sanctioning is no longer needed. The SG/E approach assumes, as does the original goal/expectation theory, that once people realize the importance of cooperation they are willing to contribute with the assurance of other members' cooperation. Thus, once other members' elementary cooperation reaches a certain level because of the sanctioning system, their fear about other members' lack of cooperation is reduced. And as a result of this reduction in fear, they are now willing to cooperate even without continued strong sanctioning. However, the lack of significant Trust × Trial Block and Sanctioning × Trust × Trial Block interaction effects suggests that this drop may not be a reliable finding. Replication studies are needed to determine whether this drop is unique to this experiment or is a more general phenomenon.

Net Benefit

The results just mentioned indicate that LTs produced a greater amount of public good (the bonus) in the sanctioning conditions than in the NS condition. However, because they paid costs for the establishment of the punishment fund, the question still remains whether they were actually better off in the HS and LS conditions than in the NS condition. In other words, did the benefit of sanctioning exceed the cost? This question was answered by comparing the net benefit (100 yen − Elementary Contribution − Instrumental Contribution + Bonus). The results of this comparison in fact indicate that LTs earned more money in the HS condition than in the NS condition in all three trial blocks: HS_1 (first trial block) = 150.6 yen per person, $NS_1 = 139.6$, $t = 2.15$, $p < .05$; HS_2 (second trial block) = 161.9, $NS_2 = 131.4$, $t = 5.03$, $p < .01$; and HS_3 (third trial block) = 179.8, $NS_3 = 129.0$, $t = 9.77$, $p < .01$ ($df = 14$. for all the t tests reported in this paragraph). HT's net benefit in the HS condition, however, was not significantly different from that in the NS condition: $HS_1 = 157.9$, $NS_1 = 160.6$, $t = -.49$; $HS_2 = 160.6$, $NS_2 = 156.5$, $t = 0.68$; and $HS_3 = 160.5$, $NS_3 = 150.0$, $t = 1.39$. On the other hand, when the effect of sanctioning was relatively weak (i.e., in the LS condition), the costs of sanctioning actually exceeded its benefits, regardless of the trust level: For LTs, $LS_1 = 91.1$, $t = -5.06$, $p < .01$, $LS_2 = 109.4$, $t = -3.86$, $p < .01$, and $LS_3 = 104.4$, $t = -3.77$, $p < .01$; for HTs, $LS_1 = 111.0$, $t = -3.89$, $p < .01$, $LS_2 = 117.4$, $t = -6.61$, $p < .01$, and $LS_3 = 105.4$, $t = -5.06$, $p < .01$. In sum, only LTs clearly benefited from the opportunity to establish the effective sanctioning system. The results just mentioned also suggest that an ineffective sanctioning system is not only useless but could actually be harmful to the total welfare of the group.

Discussion

The SG/E approach emphasizes the change in people's expectation of other members' behavior in solving the public goods problem. It assumes, as does the original goal/expecta-

tion theory, that people will cooperate once they realize the importance of cooperation and are assured of other members' cooperation. However, this change in people's expectations can be produced only through a change in the structure (such as the establishment of a sanctioning system) in large groups. In other words, the SG/E approach emphasizes the indirect effect of structural changes in addition to their direct effect. The original structural approach deals only with the direct effect of structural changes. Results of the present experiment unequivocally supported the predictions derived from this approach. Although the other two approaches taken together provide the first (without sanctioning HTs contribute more than LTs) and the third (sanctioning improves elementary cooperation) predictions, the second (LTs contribute more to the punishment fund than HTs) and the fourth (the effect of sanctioning on elementary contribution is stronger among LTs than among HTs) predictions are unique to the current approach.

Support for all four predictions demonstrate that the current approach can explain what is explained by the previous two approaches, *and more.* For example, the original structural approach can explain the higher level of elementary cooperation in the HS condition. That is, subjects contributed more in the HS condition because they anticipated higher costs for not doing so. The SG/E theory does not reject this explanation. However, the current results indicate that something additional is involved in the process besides this direct effect of sanctioning, because the direct effect of sanctioning alone cannot explain the differential effect of sanctioning among HTs and LTs. The SG/E approach suggests that this something *more* is actually the indirect effect of sanctioning via expectations of other members' cooperation. Similarly, the original goal/expectation theory can explain HT's higher level of elementary cooperation in the NS condition, but cannot explain the interaction between trust and sanctioning.

These results clearly suggest the importance of psychological factors such as trust in both elementary and instrumental cooperations. At the same time, the results suggest the importance of treating such psychological factors not simply as given but as dependent on structural factors. That is, expectations about other peoples' actions, and the fear of being exploited, can be changed by implementing a change in the structure. Finally, the results also suggest that those structural changes may be voluntarily implemented once people realize their merits and given opportunities to do so. Structural factors and psychological factors are interdependent: structural factors affect peoples' perceptions and expectations, but people act, on the basis of these perceptions and expectations, to change the structure itself.

Generalization of the current results to real public goods problems involving large groups requires careful assessment of the possible factors which distinguish large groups from small groups. The current experiment was designed to maximize compatibility with large group settings. For example, one's own elementary cooperation did not affect his or her own bonus level at all, and social incentives did not exist because subjects had no opportunity to communicate with each other. In these respects, the current experiment represents characteristics of large groups rather than those of small groups (cf. Olson, 1965). On the other hand, each individual's instrumental cooperation did actually have a tangible effect on the level of sanctioning. Because in large groups each individual's instrumental cooperation will have only a very minor effect on the level of sanctioning, the current results might not have been obtained if the group size was much larger. The SG/E approach claims, however, that whether the effect of instrumental cooperation is tangible or not does not affect members' decisions concerning instrumental cooperation, because once they adopt the goal of sustained mutual cooperation they do not act according to individual "rationality." Once they adopt this goal, the major obstacle to cooperation is not the cost-benefit factor but the fear that their cooperation may be exploited. Instrumental cooperation is much more effective than elementary cooperation for reducing this fear. Thus, people are expected to engage in instrumental but not elementary cooperation, once they adopt the goal of mutual cooperation (see Yamagishi, in press, for more discussions on this issue). However, this prediction based on the SG/E approach was not tested in the present experiment. Until this prediction is supported in future research, generalization of the current results must be tentative.

Another limitation to generalization of the current results comes from the fact that in the current experiment the chance of being "caught" for free riding was substantial. In large groups, monitoring all free riders may be difficult so that the chance of being caught for free riding may be smaller. However, the same criticism can also be directed to the original structural approach, because the validity of the structural solution is limited to situations in which sanctioning has a substantial direct effect. In this sense, the indirect effect of sanctioning predicted by the SG/E approach would be as general or as limited as the direct effect of sanctioning predicted by the original structural approach. This is actually a limitation because of effectiveness of monitoring and sanctions rather than group size per se.

Despite these limitations of the particular experimental settings used in the present study, the merits of the general approach proposed here are clear. In past research on social dilemmas, people have often been regarded as simply responding to an environment that is beyond their control (cf. Buckley, Burns, & Meeker, 1974; Edney & Harper, 1978). The current approach, on the other hand, emphasizes that people not only passively respond to their environment but also actively create, maintain, and change it. The results of the current study demonstrate the importance of this perspective in studying social dilemmas. Second, the current study demonstrates the fruitfulness of analyzing the interaction between structural factors and psychological factors. The analysis of the interaction between these two factors provides both scholars of large scale public goods problem and experimental researchers of social dilemmas an opportunity to recognize each other's contributions and to enrich their own work by incorporating the contributions made by the other type of research. Finally, the current study points out the importance of higher order (second-order, third-order, etc.) public goods, and thus it will hopefully provide a stimulus for more discussion and research on the unique characteristics of instrumental as against elementary cooperation.

References

Acheson, J. M. (1975). The lobster fiefs: Economic and ecological effects of territoriality in the Maine lobster industry. *Human Ecology, 3,* 183–207.

Alcock, J. E., & Mansell, D. (1977). Predisposition and behavior in a collective dilemma. *Journal of Conflict Resolution, 21,* 443–457.

Bonacich, P., Shure, G. H., Kahan, J. P., & Meeker, R. J. (1976). Cooperation among six persons in a prisoner's dilemma. *Journal of Conflict Resolution, 20,* 687–706.

Brubaker, E. R. (1975). Free ride, free revelation, or golden rule? *Journal of Law and Economics, 18,* 147–161.

Buckley, W., Burns, T., & Meeker, L. D. (1974). Structural resolutions of collective action problems. *Behavioral Science, 19,* 277–297.

Caldwell, M. D. (1976). Communication and sex effects in a five-person prisoner's dilemma. *Journal of Personality and Social Psychology, 33,* 273–281.

Cass, R. C., & Edney, J. J. (1978). The commons dilemma: A simulation testing the effects of resource visibility and territorial division. *Human Ecology, 6,* 371–386.

Edney, J. J., & Harper, C. S. (1978). The commons dilemma: A review of contributions from psychology. *Environmental Management, 2,* 491–507.

Fox, J., & Guyer, M. (1977). Group size and others' strategy in an *N*-person game. *Journal of Conflict Resolution, 21,* 323–338.

Goehring, D. J., & Kahan, J. P. (1976). The uniform *n*-person prisoner's dilemma game. *Journal of Conflict Resolution, 20,* 111–128.

Hamburger, H., Guyer, M., & Fox, J. (1975). Group size and cooperation. *Journal of Conflict Resolution, 19,* 503–531.

Kelley, H. H., & Grzelak, J. (1972). Conflict between individual and common interest in an *n*-person relationship. *Journal of Personality and Social Psychology, 21,* 190–197.

Komorita, S. S., & Lapworth, C. W. (1982). Cooperative choices in social dilemmas. *Journal of Conflict Resolution, 26,* 692–708.

Komorita, S. S., Sweeney, J., & Kravitz, D. A. (1980). Cooperative choice in the *N*-person dilemma situation. *Journal of Personality and Social Psychology, 38,* 504–516.

Linaweaver, F. P., Geyer, J. C., & Wolff, J. B. (1967). Summary report on the residential water use research project. *Journal of American Water Works Association, 59,* 267–282.

Maki, J. E., Hoffman, D. M., & Berk, R. A. (1978). A time series analysis of the impact of a water conservation campaign. *Evaluation Quarterly, 2,* 107–118.

Marwell, G., & Ames, R. E. (1979). Experiments on the provision of public goods, I: Resources, interest, group size, and the free-rider problem. *American Journal of Sociology, 84,* 1335–1360.

Marwell, G., & Ames, R. E. (1980). Experiments on the provision of public goods, II: Provision points, experience, and the free-rider problem. *American Journal of Sociology, 85,* 926–937.

Marwell, G., & Schmitt, R. (1972). Cooperation in a three-person prisoner's dilemma. *Journal of Personality and Social Psychology, 21,* 376–383.

Messick, D. M., & McClelland, C. (1983). Social traps and temporal traps. *Personality and Social Psychology Bulletin, 9,* 105–110.

Messick, D. M., Wilke, H., Brewer, M. B., Kramer, R. M., Zemke, P. E., & Lui, L. (1983). Individual adaptations and structural change as solutions to social dilemmas. *Journal of Personality and Social Psychology, 44,* 294–309.

Oliver, P. (1980). Rewards and punishments as selective incentives for collective action: Theoretical investigations. *American Journal of Sociology, 85,* 1356–1375.

Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups.* Cambridge, MA: Harvard University Press.

Pruitt, D. G., & Kimmel, M. (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *American Review of Psychology, 28,* 363–392.

Samuelson, C. D., Messick, D. M., Rute, C. G., Wilke, H. (1984). Individual and structural solutions to resource dilemmas in two cultures. *Journal of Personality and Social Psychology, 47,* 94–104.

Sato, K. (1984). *Effects of the allocation of negative common goods on group productivity.* Paper presented at the International Congress of Psychology, Acapulco, Mexico.

Sato, K. (1985). *Distribution of the cost for maintaining common resources.* Paper submitted for publication. (Mimeo, Department of Behavioral Science, Hokkaido University, N10 W7 Kita-ku, Sapporo, Japan 060).

Sato, K., & Yamagishi, T. (in press). Two psychological factors in the problem of public goods. *Japanese Journal of Experimental Social Psychology.* (In Japanese)

Yamagishi, T. (in press). The structural goal/expectation theory of cooperation in social dilemmas. In E. J. Lawler & B. Markovsky (Eds.), *Advances in group processes.* Vol. III. Greenwich, CT: JAI Press.

Yamagishi, T., & Sato, K. (1986). Motivational basis of the public goods problem. *Journal of Personality and Social Psychology, 50,* 67–73.

Yukawa, M. (1985). *Structural and Psychological Factors in Social Dilemmas.* Unpublished bachelor's thesis, Hokkaido University, Sapporo, Japan.

Received August 21, 1985
Revision received December 19, 1985 ■