Organizing Groups for Collective Action

Robyn M. Dawes; John M. Orbell; Randy T. Simmons; Alphons J. C. Van De Kragt

The American Political Science Review, Vol. 80, No. 4. (Dec., 1986), pp. 1171-1185.

ORGANIZING GROUPS FOR COLLECTIVE ACTION

ROBYN M. DAWES
Carnegie-Mellon University
JOHN M. ORBELL
University of Oregon and *Utah State University*
RANDY T. SIMMONS
Utah State University
ALPHONS J. C. VAN DE KRAGT
Carnegie-Mellon University

How can the beneficiaries of collective action be persuaded to contribute the resources (time, energy, money) necessary for the effort to succeed? Rational and selfish players will recognize they can free ride on the successful contributions of others. If the effort is not successful, they will lose a contribution—and be "suckered." Other than relying on altruism, organizers of the group effort can modify incentives so that players are more prepared to contribute. Laboratory experiments offer one way of assessing the effectiveness of various such modifications; we conducted such tests to see how well contributing is promoted by (1) assuring contributors that they will not lose if the group effort fails (a "money-back guarantee") and (2) enforcing contributions if it succeeds ("fair share"). We expect the latter to be more successful because it is "stable," unlike the former, whose success can be undermined by expectations of that success. Three experimental replications demonstrate that the money-back guarantee is no more successful than a standard dilemma, but fair-share requirements increase contributing significantly over that base. Analysis of subjects' expectations about others' behavior offers some support to the hypothesized process undermining the money-back guarantee, but motivational factors must also be taken into account for a full explanation.

The problem of organizing groups for collective action (whether political or otherwise) has, since Olson (1965), been well understood: at least when the group is large, there is a clear disincentive for the potential beneficiaries to contribute the time, money, or other resources necessary for the group effort to be successful, even if all want that effort to be successful. Why should *I* contribute when there is only a trivial chance that my contribution will make a critical difference and when the only other possibilities are that the group effort will be realized without my contribution (in which case I can "free ride") and that it will not be realized if I do contribute (in which case I will be "suckered")?

In more formal terms, the problem is a *social dilemma* (Dawes, 1980): it involves a dominant incentive (do not contribute) associated with a suboptimal equilibrium (the group effort fails). If individuals follow their self-interest, groups confronting such incentives will not attain objectives that all members want.

In natural situations, of course, individuals do not always appear to follow their self-interest, and this impression is supported by experimental studies of social dilemma behavior in which it is possible to specify payoffs with confidence; some incidence of cooperation is normal in such experiments, much more if a period of group discussion is permitted (see, for example, Caldwell, 1976; Dawes, McTavish, and Shaklee, 1977; Edney and Harper, 1978; Jerdee and Rosen, 1974; Orbell, Schwartz-Shea, and Simmons, 1984; Rapoport, 1974; van de Kragt, Orbell, and Dawes, with Braver and Wilson, 1986). To the extent people do not respond to self-interest in these games, the pessimistic prediction from a priori analysis is not borne out, and the side payments Olson discusses (whether "threats" or "bribes") are not required.

In spite of these data—indeed, because the incidence of cooperation (in one-shot games) in the absence of discussion is, characteristically, low—good sense dictates that we search for institutional ways of solving social dilemmas; that is, that we restructure incentives so that selfish individuals are led by consideration of their private interests to contribute to the common interest.[1] Such a restructuring is in the spirit of Adam Smith's (1976) "unseen hand" and of the derived modern arguments in favor of fully specified private property rights and the market as a mechanism for making social decisions. It is also in the spirit of America's founding fathers and of contemporary efforts to design institutions for collective decision making so that selfish individuals playing political games are led to act in the common interest.

Restructuring incentives is what we investigate in the present article. We offer an empirical study of the relative effectiveness of two ways of modifying incentives in step-level dilemmas: (1) ensuring players that they will not lose a contribution if the group effort fails and (2) enforcing a contribution (requiring a "fair share") in the event the group effort succeeds.

The *money back guarantee* device was used by a group of state system faculty members (The Association of Oregon Faculties) wishing to raise money in 1979 to hire a lobbyist at the state legislature. The desired lobbyist—a public good for all faculty members because any pay increases he produced would go to all faculty in the system, not just to those who contributed—required an annual retainer of $30,000 for his work.[2] The Association asked all faculty in the state for contributions, suggesting $36, $60, or $84, depending on salary. The request contained an explicit promise that all money would be returned if less than the $30,000 was raised. The solicitation was successful.

The logic of *enforced contribution* is often present in efforts by apartment dwellers to resist developers who wish to convert their apartment building into a condominium. The developers offer to sell the units at a reduced rate to anyone wishing to vote for conversion prior to a specified deadline. One "contributes" to the apartment dwellers' effort, then, by withstanding the offer. If the effort fails and the conversion proceeds, those who withstood the offer are out of pocket to the extent of the reduced offer; they have to pay the higher rate. However, it is not possible for an apartment dweller to free ride on the restraint of others because if a sufficient number withstand the offer, the conversion won't occur, and the people who voted "for" won't get the private benefit from the sale. Similar logic can exist in union elections in which the choice is between supporting or not supporting an organizing effort and where, if the vote succeeds, "fair share" provisions exist and all members of a bargaining unit must pay their fees.

In natural circumstances, a wide range of uncontrolled factors can influence outcomes one way or the other, and we cannot say—short of the most detailed study—whether success of the group effort hinges on the presence (or absence) of one or the other of these devices. However, the comparative usefulness of the respective organizational devices is an issue susceptible to theoretical and experimental investigation, and the first purpose of the present study is to compare the operation of the two in an experimental context in which factors that might influence performance in natural circumstances are random. We address this purpose in Experiments 1 and 2.

Our design involves a step-level game (Frohlich and Oppenheimer, 1970; Rapoport, 1985; van de Kragt et al., 1983) in which a "bonus," or public good, is provided to all group members in the event some specified number of them make a fixed contribution. In the *full dilemma*—the standard version of this game—individuals have a positive incentive for not contributing if enough others do contribute *or* if too few others contribute to make a contribution useful. In the former case, the individual can "free ride," and in the latter he or she can avoid wasting a contribution. Our institutional modifications involve *half dilemmas* in which one of these incentives is removed. These are the *money-back guarantee half dilemma* and the *enforced contribution* ("fair share") *half dilemma*, respectively. With the money-back guarantee, individuals receive higher payoffs for not contributing if the public good is provided and no more than contributors if it is not. With the enforced contribution, they receive higher payoffs for not contributing if the public good is not provided and no more than contributors if it is.

There is a third possible outcome of others' choices in any step-level dilemma, whether full or half: that the individual's contribution will be "critical," or necessary for provision of the public good; that is, when k contributions are required, $(k - 1)$ others will contribute and the individual's contribution will make the difference.[3] Setting aside this possibility, we note that players in the half dilemmas have a dominating incentive to withhold a contribution, just as do players in the standard game: they will make more by defecting if one outcome happens and no less by defecting if the other does.

At least in the large number case, it does seem reasonable to set aside the possibility of "criticalness." It is, to use our examples, highly unlikely that all (or, indeed, any) contributing faculty or apartment dwellers contributed because they believed their contribution would be critical. Yet this *is* possible, especially in smaller groups, and it is also possible that there are different beliefs about the probability of being critical in the three conditions. A second purpose of this set of experiments, therefore, is to examine the effects of the money-back guarantee and the enforced contribution on the subjective probability of (1) being critical, (2) wasting one's contribution, and (3) being redundant. We address this purpose in Experiment 3, in which we collect such probability judgments. In addition, we will use these probabilities to estimate the expected value of the payoffs for contribution or noncontribution in the standard dilemma and the two half dilemmas. These expected values are not the same in the three conditions, and we will attempt to relate actual choice between contributing and not contributing to these values both within and across conditions.

Predictions

As Coombs (1973) has pointed out, the prisoner's dilemma offers two incentives to defect: the desire to avoid being "suckered" and the desire to capture the "free rider's" payoff. In Coombs' terms, these incentives are *fear* and *greed*, respectively, and they are redundant; either is sufficient to predict defection by itself. We note that our two half dilemmas each retain one of these incentives while eliminating the other. Providing a money-back guarantee offers the individual protection against loss through contributing, while enforcing a contribution should the good be provided means that nothing is to be gained from withholding a contribution—if the good is provided under enforced contribution, a player will end up paying one way or the other. In 1975, Brubaker proposed that defection was more a consequence of desire to avoid loss through contributing than of a desire for gain. If he is correct, we would expect greater group success from the half dilemma that offers a money-back guarantee than from the half dilemma that enforces contribution.

However, a structural difference between the two predicts the opposite. Let us suppose that people given a money-back guarantee believe that such a guarantee works—that is, it will increase the probability that others will contribute. This belief yields an increase in the subjective probability of the goods being obtained without the individual's contribution and, therefore, an increase in the probability that an attempt to free ride will be successful. For example, a faculty member who registered the money-back guarantee might have assumed that its presence would induce enough other faculty members to contribute to make his or her own contribution unnecessary. In short, the money-back guarantee may not "reduce" the dilemma at all. When such a guarantee is given, people may respond to it no differently from the way they respond to the standard dilemma and contribute for the same reasons they contribute in that standard situation (presumably conscience, fellow feelings, magnanimity, etc.) and in the same numbers. *The instability of the money-back guarantee half dilemma may result in its collapse as a device for improving base rates of contributing.*[4]

In contrast, the enforced contribution dilemma is stable. Suppose that people who know they are operating under an enforced contribution rule believe that it will work—that is, it will increase the probability that others will contribute. This belief implies a decrease in the subjective probability that the good will not be provided by others' contributions—it makes group failure and the attendant reason for not contributing less likely. Why *not* contribute under such circumstances?

From this nonmotivational analysis, we predict that the enforced contribution half dilemma will be more effective in raising contributions than will the money-back guarantee half dilemma.

Methodological Overview

Predictions such as these can be readily tested in the laboratory where all other factors that might influence cooperation rates in natural situations are randomized. We do not claim that public goods games created in the laboratory are simulations of any particular naturally occurring situation; we claim only that the payoff matrices confronting subjects *are* public goods matrices. Similarly, we are not particularly interested in absolute levels of cooperation, which can be expected to vary with subject pools and (as we will see) historical periods. Our interest is in differences in cooperation that might exist between experimental conditions of theoretical interest, and we know of no good reasons for predicting an interaction between condition and population pool.

We created seven-person public goods problems under such laboratory circumstances. After they signed the necessary release form, subjects were seated around a large table and were read instructions.

Each was given a $5 promissory note and was told that he or she could contribute that $5 toward a "bonus" of $70 to the group as a whole. This bonus would be distributed equally so that each group member would receive $10. A specified minimal number of contributions was required for the bonus to be distributed; if that number of contributions was not forthcoming, there would be no bonus. Subjects made the decision between contributing and not contributing simultaneously, anonymously, and only once. The latter fact meant there was no possibility of influencing other subjects' behavior in the future, a consideration crucial to some proposed solutions to social dilemma games—notably Axelrod and Hamilton's (1981). Each subject was paid singly and was on the elevator before the next one was paid; this procedure was explained to the subjects in advance so that they knew their choices would be completely unknown to the other participants. If the subject had not contributed the $5 and the bonus was produced, then the net payoff was $15; if the subject had contributed and the bonus was produced, the net payoff was $10; if the subject had not contributed and not enough other subjects had contributed to produce the bonus, the net payoff was $5; and if the subject had contributed and not enough others had, the net payoff was zero. As was emphasized at the outset, there was no deception. Subjects' payoffs varied, as indicated, as a function of their own choices and of the simultaneous choices of the others in the experimental session.

The game described above is a standard step-level public goods game, one that involves no communication, no opportunity for persuasion or coercion, no possibility of side payments or reciprocity, and no social disclosure of individual choices (except to the experimenters). We modified it in two ways: the money-back guarantee half dilemma and the enforced contribution half dilemma.

Under the money-back guarantee half dilemma, subjects were told that if they contributed their $5, but there were not enough other contributions to ensure provision of the bonus, their $5 would be returned. Thus, there was no way in which subjects could lose their contributions. However, they could still free ride on the contributions of others. If a subject withheld his or her contribution and enough others contributed to provide the bonus, the subject would benefit from both the bonus and the original $5, obtaining a total of $15 for participating in the experiment.

Under the enforced contribution half dilemma, all payments were truncated at $10. Should the bonus be provided, no subject could leave the experiment with more than $10. (Truncating payments at $10 is logically equivalent to forcing all to contribute $5 if the bonus is provided.) Thus, there was no opportunity to free ride. When the bonus was provided by others' contributions, subjects who did not contribute left with the same amount of money—$10—as those who had contributed. However, subjects could still lose their contributions and be "suckered." If enough other people did not contribute to produce the bonus, then individuals who did contribute would lose their $5.

The main difference between our experimental dilemmas and the naturally occurring ones discussed earlier is size; with the smaller numbers demanded by experimental logistics, the probability of being critical may be seen as important. Such beliefs will be discussed, evaluated, and analyzed in Experiment 3. The purpose of these first two experiments is to compare the relative efficiency of the two half dilemmas in enhancing cooperation and to test the prediction that enforcing a contribution in the event the good is provided will elicit more voluntary contributing than will providing a money-back guarantee.

Table 1. Experiment 1: Percentage Contributing, Number of Contributors, and the Analysis of Variance and Scheffé Test Results (3 or 7 required)

		Standard Dilemma (1)	Money-Back Guarantee (2)	Enforced Contribution (3)
Percentage contributing		51	61	86
Number of contributors in each group		1, 2, 2, 3, 3, 3, 4, 5, 6, 7	3, 3, 4, 4, 5, 5, 6	4, 5, 5, 6, 6, 6, 7, 7, 7, 7
Analysis of Variance	*df*	*MS*	*F*	*p*
Between conditions	2	14.05	7.41	.003
Within conditions	24	1.90		
Total	26	2.83		
Post hoc Scheffé tests[a]	Average difference	Range for $p = .10$	Range for $p = .05$	Range for $p = .01$
(1) vs. (2)	.59	1.52	1.76	2.32
(1) vs. (3)	2.30	1.38	1.59	2.04
(2) vs. (3)	1.71	1.52	1.76	2.32

[a]Any difference between means larger than the number specified is significant at the level indicated.

Experiment 1

Subjects were recruited by advertising in the Eugene, Oregon, *Register-Guard* and in student newspapers at the University of Oregon and Utah State University. Subjects participated in the standard dilemma game, the money-back guarantee half dilemma, or the enforced contribution half dilemma. As much care as possible was taken to ensure that subjects in a particular session did not know each other prior to the experiment. In each condition, contributions from any three or more subjects in seven-subject groups were required for the bonus to be provided. The standard dilemma game and the money-back guarantee half dilemma were run in 1979 (forming the core of Simmons's doctoral dissertation; see Simmons, 1980), while the enforced contribution half dilemma was conducted two years later.[5]

All subjects were given the $5 promissory note, after which the appropriate experimental condition was explained in detail. The explanation covered both the logic of the game and the explicit set of monetary payoffs under each condition. Subjects were "quizzed" about the payoffs, and if there was any misunderstanding, the instructions were explained again. The individual payoffs and anonymity were emphasized. *Subjects were not permitted to communicate among themselves; any questions were directed to the experimenter.* In all conditions, subjects made their choices—either to withhold the $5 promissory note or to contribute it —by checking the appropriate box on the "decision form" they had in front of them.

Results of Experiment 1

Table 1 presents the percentage of subjects contributing, the number of contributors, the analysis of variance results, and the results of a *post hoc* Scheffé test using groups as the unit of analysis.[6] The results indicate that there is no significant difference between contribution rates in the standard dilemma and the money-

Table 2. Experiment 2: Percentage Contributing, Number of Contributors, and the Analysis of Variance and Scheffé Test Results (5 of 7 required)

		Standard Dilemma (1)	Money-Back Guarantee (2)	Enforced Contribution (3)
Percentage contributing		64	65	93
Number of contributors in each group		3, 4, 4, 4, 4, 4, 5, 5, 6, 6	3, 4, 4, 5, 5, 5, 6	5, 6, 6, 6, 7, 7, 7, 7, 7, 7
Analysis of Variance				
	df	*MS*	*F*	*p*
Between conditions	2	12.24	15.69	.00
Within conditions	24	.78		
Total	26	1.66		
Post hoc Scheffé tests[a]				
	Average difference	Range for $p = .10$	Range for $p = .05$	Range for $p = .01$
(1) vs. (2)	.07	.99	1.14	1.46
(1) vs. (3)	2.00	.89	1.03	1.32
(2) vs. (3)	1.93	.99	1.14	1.46

[a]Any difference between means larger than the number specified is significant at the level indicated.

back guarantee half dilemma (51% and 61%, respectively) while there is a significant difference between contribution rates in the standard dilemma and the enforced contribution half dilemma (51% vs. 86%).

Experiment 2

This experiment was identical in all respects to Experiment 1, except that five out of the seven subjects were required to make a contribution in order for the bonus to be provided. By requiring five contributions instead of three, questions of criticalness, deficiency, and redundancy were reversed. For example, whereas two of six others had to contribute for an individual's contribution to be critical in Experiment 1, four of six others in this experiment had to do so.[7]

Results of Experiment 2

The outcomes of Experiment 2 are given in Table 2. Once again, there is no significant difference between contribution rates in the standard dilemma and the money-back guarantee half dilemma (64% and 65%, respectively) while there is a significant difference between contribution rates in the standard dilemma and the enforced contribution half dilemma (64% vs. 93%). Note that the contribution rates are somewhat higher in this experiment (where five contributions are required to activate the bonus) than in the first experiment (where three contributions are required). An analysis of variance in which the main effect is three vs. five required contributions shows that the number required does not produce a significant main effect ($F_{1,48} = 3.32$, $p >$ .07) and that its interaction with the three conditions is not significant ($F_{2,48} = .196$, $p > .7$).

As predicted by the structural hypothesis, the enforced contribution half dilemma was superior to the money-back guarantee. In fact, the contribution rate in the money-back guarantee half dilemma was not significantly higher than the rate in the standard dilemma in either experiment while the differences between the enforced contribution half dilemma and the standard dilemma were large and highly reliable. Whether the money-back guarantee works at all is a moot point; across both experiments it yields a cooperation rate of 63%, as opposed to 57% in the standard dilemma. (A rough median split indicates that if these percentages remained constant, the results would be significant if the sample were multiplied by a factor of about 10.)

Experiment 3

What accounts for the 60% cooperation rate in our standard dilemma without any structural modification? One possible explanation is that such base-line contributors were influenced by motivations unrelated to their own monetary payoff—for example, conscience, altruism, and self-esteem. Another explanation, however, derives from the step-level character of our basic design. These contributors might have believed the probability that they were critical to the provision of the bonus (to themselves, as well as to the group) was greater than .5, in which case they would receive a net gain of $5 by contributing, not a net loss of $5, as they would under any other circumstance. The purpose of our third experiment was to distinguish between these two hypotheses by investigating the relationship between subjective probability estimates about others' contributions and a person's own choice, both within and between groups. This approach was suggested by Rapoport (1985), who begins by postulating that in each of these experimental goods situations, each participant estimates three probabilities:[8]

1. p = the probability of being futile—that is, the probability that whatever one does, an insufficient number of others will contribute for the public good to be provided (less than $[m - 1]$ others when m or more are required);
2. p^* = the probability of being critical—that is, the number of others contributing is exactly the number such that the bonus is provided if and only if the individual contributes (exactly $[m - 1]$ others when m or more are required);
3. $(1 - p - p^*)$ = the probability of being redundant—that is, a sufficient number of others will contribute to ensure that the public good will be provided regardless of what the individual chooses (m or more others when m or more are required).

Rapoport works out the expected values for contributing and not contributing. For example, consider the value of cooperation to be $V(c)$ and that of the bonus to be $V(g)$. Letting the subscript 1 refer to the full dilemma condition, we compute the expected value of contributing to be

$$0 \times p_1 + V(g) \times (1 - p_1) = V(g) \times (1 - p_1). \quad (1)$$

That of not contributing is

$$(p_1 + p^*_1) \times V(c) + (1 - p_1 - p^*_1) \times [V(g) + V(c)] = V(c) + (1 - p_1 - p^*_1) \times V(g). \quad (2)$$

The difference is $[p^*_1 \times V(g) - V(c)]$, which is greater than zero only if

$$p^*_1 > V(c)/V(g) = 1/c, \quad (3)$$

where $[c = V(g)/V(c)]$.

Hence, an individual maximizing expected value would choose to contribute if and only if inequality (3) were satisfied. Similar algebra demonstrates that in the money-back guarantee half dilemma, the expected value of contributing is greater than that of not contributing if and only if

$$p^*_2 > (1 - p_2)/c, \quad (4)$$

where the subscript 2 indicates probability estimates for this second situation. Finally, it can be demonstrated that for the third situation—the enforced contribution half dilemma—the expected value of contributing is greater than that of not contributing if and only if

$$p^*_3 > p_3/(c - 1). \quad (5)$$

Rapoport derives the same three conditions. Note first that all three involve an inequality implying that the maximizer of expected value should contribute if and only if p^* is greater than some value. The inequality involves p^* because the situation in which the potential contributor is critical is the only one in which he or she experiences a net gain of $5 by contributing, rather than a net loss of $5 or indifference. Note also that all three inequalities involve only probabilities and the ratio of $V(c)$ and $V(g)$, that is, c, which in our experiments is equal to two. These equations allow qualitative predictions, even though the actual probabilities p, p^*, and $(1 - p - p^*)$ may vary by condition. That is, within each condition we can compare the choice of each individual to his or her estimate of p^* and in conditions 2 and 3, of p.

Method

We replicated Experiment 2 (5 or more contributions out of 7 required to elicit the bonus) with subjects recruited by advertisements in the student newspaper at Utah State University, Logan, Utah, in late November and early December 1984. In addition to asking subjects to make a choice, we asked them each to estimate the probabilities p, p^* and $(1 - p - p^*)$. As before, each subject was assigned to only one condition on a quasi-random basis, and care was taken that friends or acquaintances were not assigned to the same group. The instructions were identical to those of Experiment 2. The probability estimates were elicited by asking subjects to distribute 100 points, as follows:

> For the following questions you must enter a number between 100 and zero. The number indicates how likely you think it is that the specified outcome will occur. These numbers must total to 100. If you are certain that a specific outcome will occur, put 100 in that box and 0 in the others. If you think that each outcome is equally likely, put 33.3 in each box. Any combination is possible, of course, but remember that all numbers must add to 100.

We then asked for answers to the following questions:

1. What is the likelihood of fewer than four others choosing to invest, that is to say, one, two or three other members of the experiment?[9]
2. What is the likelihood of exactly four others choosing to invest?
3. What is the likelihood of more than four others choosing to invest, that is to say, five or all six of the other members of the experiment?

Five groups were run in each of the conditions: standard dilemma, money-back guarantee half dilemma, and enforced contribution half dilemma.

Results of Experiment 3

Table 3 presents the percentage contributing, the number of contributors in each group, and the analysis of variance and a *post hoc* Scheffé test result between conditions. Note that the overall contribution rate is less in 1984 than in 1979–81, significantly so ($F_{1,40} = 15.14$, $p < .0005$). Nevertheless, the pattern is the same: there is no significant difference between the standard dilemma and the money-back guarantee half dilemma, and there is a significant difference between

Table 3. Experiment 3: Percentage Contributing, Number of Contributors, and the Analysis of Variance and Scheffé Test Results (5 of 7 required)

		Standard Dilemma (1)	Money-Back Guarantee (2)	Enforced Contribution (3)
Percentage contributing		23	43	77
Number of contributors in each group		1, 1, 2, 2, 3	5, 2, 1, 3, 4	4, 5, 6, 6, 6
Analysis of Variance				
	df	*MS*	*F*	*p*
Between conditions	2	18.47	15.39	.0005
Within conditions	12	1.20		
Total	14	3.67		
Post hoc Scheffé tests[a]				
	Average difference	Range for $p = .10$	Range for $p = .05$	Range for $p = .01$
(1) vs. (2)	1.40	1.64	1.93	2.58
(1) vs. (3)	3.80	1.64	1.93	2.58
(2) vs. (3)	2.40	1.64	1.93	2.58

[a]Any difference between means larger than the number specified is significant at the level indicated.

the full dilemma and the enforced contribution half dilemma.

As we pointed out earlier, our interest is in behavioral differences between theoretically relevant conditions in experiments such as ours not in absolute levels of cooperation, which can be expected to reflect cultural and other parametric differences. It is easy—no doubt too easy—to speculate on the causes, as well as the consequences, of what here seems to be a behavioral difference between historical periods.

Table 4 gives the average probability estimates, along with their standard deviations, of being futile, being critical, and being redundant in the three conditions. The estimates are broken down into cooperators vs. defectors. A three-way repeated measures analysis of variance on the data of the second part of Table 4 (treating the estimated probabilities of being futile and of being redundant for each individual nested within conditions as the repeated variable) reveals that probability estimates vary by condition ($F_{2,102} = 5.35$, $p = .01$).[10]

The pattern is clear. The probability of redundancy is enhanced and the probability of deficiency is diminished in both half dilemmas by a ratio of 2:1. The judged probability of being critical is unaffected. This pattern of average estimates is consistent with the expectations model of cooperation; it follows that *on the average*, a subject should cooperate in the three conditions if p^* is greater than .50 for equation (3), .37 for equation (4), and .27 for equation (5). These inequalities are progressively easier to satisfy—but note empirically that, on the average again, none is itself easy to satisfy. Note also that, since the p^*s are inherently incomparable, this prediction is "weak."

Moreover, the analysis shows that there is a main effect for cooperation ($F_{1,102} = 26.36$, $p < .001$) with cooperators having higher estimates of being critical or redundant than defectors. More importantly, there is an interaction effect

Table 4. Experiment 3: Average Probability Estimates, with Standard Deviations, of Being Futile, Critical, and Redundant, and Analysis of Variance for Repeated Measures

Experimental Condition and Choice	Probability of Being Futile ($M < 4$) Average Probability Estimate	Probability of Being Futile ($M < 4$) Standard Deviation	Probability of Being Critical ($M = 4$) Average Probability Estimate	Probability of Being Critical ($M = 4$) Standard Deviation	Probability of Being Redundant ($M > 4$) Average Probability Estimate	Probability of Being Redundant ($M > 4$) Standard Deviation
By condition						
Standard dilemma	.47	.27	.21	.20	.32	.25
Money-back guarantee	.27	.17	.25	.18	.48	.23
Enforced contribution	.27	.22	.24	.15	.50	.26
By choice and condition						
Standard dilemma						
Cooperators ($N = 9$)	.24	.20	.29	.31	.46	.32
Defectors ($N = 26$)	.55	.25	.18	.13	.26	.20
Money-back guarantee						
Cooperators ($N = 15$)	.25	.15	.29	.19	.46	.18
Defectors ($N = 20$)	.28	.19	.21	.16	.50	.28
Enforced contribution						
Cooperators ($N = 27$)	.20	.16	.23	.15	.57	.25
Defectors ($N = 8$)	.51	.23	.25	.16	.24	.08

for cooperation by condition ($F_{2,101} = 5.91$, $p < .001$). Thus, the estimates of cooperators and defectors are differentially affected by the changes from the standard condition to the two half dilemmas. Are these changes compatible, however, with the hypothesis that cooperation or defection are due to differential expectations aroused by the different conditions?

First, the actual numbers do not support the hypothesis. It is specifically the defectors whose expectations should be compatible with not contributing, but under the standard condition, the defectors' average probability estimates of either deficiency or redundancy is .81, as opposed to .70 for the cooperators (who—according to the model—should not cooperate in such circumstances). Under the money-back guarantee condition, the defectors' average probability estimate of being redundant (which makes free riding possible) is virtually identical to that of the cooperators: .50 vs. .46. Only in the enforced contribution condition are the differential directional predictions for cooperators as a group vs. defectors as a group supported. The average probability estimate of the defectors (there were only eight) of being deficient is .51, as opposed to .20 for the cooperators.

When the predicted behavior of each separate individual is compared with his or her actual behavior, directional results are more compatible. Here we computed the phi value of the 2 × 2 contingency between whether the subject did or did not contribute and whether, for that subject, inequality (3), (4), or (5) is satisfied in the appropriate condition. In the standard dilemma, 7 of the 9 contributors are predicted defectors while none of the defectors is a predicted contributor; phi is .42. Under the money-back guarantee condition, 12 of the 15 contributors are predicted defectors and, of the 20 defectors, 3 are predicted cooperators; phi is .09. Under the enforced contribution condition, 13 of the 27 contributors are predicted defectors while 2 of the 8 defectors are predicted contributors; phi is .23. Pooling across the three conditions, only 19 of the 51 contributors "should have" contributed, while 5 of the 54 defectors "should have" done so too; phi is .33 (chi-square = 11.52, $df = 3$, $p < .01$). Thus, the direction of contingency supports the predictions of expectations, but note that a full 63% of the contributors had a negative expected value for contributing. The phi coefficient is, however, much larger in the standard and enforced contribution conditions than in the money-back guarantee, a finding that could follow from the hypothesized instability of the money-back guarantee condition.

A final individual analysis involved the contingency between contributing or not and estimating that such a contribution would result in a net loss of $5 relative to potential payoff. In the standard condition, such a $5 loss occurs if an individual's contribution would be either deficient or redundant. Of the 9 cooperators, 7 believed that they had a better than .50 probability of sustaining such a loss—that is, the sum of each person's estimated probabilities of deficiency and redundancy was greater than .50. In contrast, 25 of the 26 defectors believed that they had a probability of over .50 of sustaining such a loss. Again, the absolute numbers do not support the expectations model, but the direction does, although not significantly so; phi = .29.

Under the money-back guarantee condition, such a $5 net loss occurs only if the individual is redundant—if more than four others also contribute. Of the 15 cooperators, 5 believed that they had a better than .50 chance of sustaining such a loss while 8 of the 20 defectors believed that they did. Here, even the directional contingency is virtually nonexistent; phi = .07.

Under the enforced contribution condition, such a $5 net loss to a contributing individual occurs only if that individual is deficient—if fewer than four others also contribute. Only 1 of 27 cooperators believed that he or she had a better than .50 probability of sustaining such a loss, while 3 of 8 defectors believed that they did. Once more, the absolute numbers do not support the expectations model, but the direction does; phi = .45. We also compared the last two contingencies by means of a Goodman z test for analyzing the interaction term in a 2 × 2 × 2 table.[11] The resulting value was 1.71 ($p < .05$, one-tailed), but this value must be treated as merely suggestive, due to the *post hoc* nature of the test and because two cells had expected values of less than 5.0 (1.93 and 3.63).

Finally, pooling across all conditions, of 49 subjects who expected a $5 relative loss with probability greater than .50 were they to contribute, only 13 did. In contrast, of 56 who did not estimate the probability of such a relative loss to be greater than .50, 38 cooperated. This contingency is not only strong (phi = .41), but highly significant; chi-square = 17.87 with 3 degrees of freedom, $p < .001$.

We must point out that the analysis based on potential net loss ignores the outcome in which contributing or not leads to the same outcome.[12] Nevertheless, the results of the analysis constitute a clear comparison of the money-back guarantee with the enforced contribution. The enforced contribution results are more compatible with the expectations model than are the money-back guarantee results (again supporting the hypothesis of instability of expectations under the money-back guarantee). Moreover, in the enforced contribution, the contingency between estimation and choice is stronger when the possibility that leads to the same result for contributing and not contributing is ignored.

Discussion

There is no ambiguity whatever about the success of the money-back guarantee device for eliciting contributions compared with the success of the enforced contribution device: the enforced contribution is superior. The pattern of results is stable over variations in the rate of contribution between regions and over a five-year time span. In all three replications it was superior to the standard dilemma at a high level of statistical significance. In two of three, it was significantly superior to the money-back guarantee and was marginally better in the third. In contrast, in no experiment did the money-back guarantee yield results that were superior to those in the standard dilemma at standard levels of significance.

Much of what social science can demonstrate is already "known" by evolved social systems, and we speculate that the frequent use of the enforced contribution or "fair share" device—most notably, of course, in labor unions—is a consequence of long experience with alternative ways of trying to persuade the beneficiaries of group efforts to contribute to those efforts. Institutions that work or that add nothing beyond some base tend to be found out, as do institutions that do not work.

Yet if frequent use suggests that a mechanism has been found to work, it does not, in itself, explain why that mechanism works. We proposed that "fair share," or enforcing a contribution in the event the group effort succeeds, is successful because expectations of its success do not undermine that success—in contrast to the money-back guarantee. However, while the behavioral data supported the prediction, the expectations data (Experiment 3) lend only directional support. There is clearly a contingency between choice and expectation based on probability estimates, but the actual behavior must be influenced by other factors as well. We speculate that one such factor is the relative motivational strengths of Coombs's (1973) *fear* and *greed*.

Earlier, we mentioned Brubaker's (1975) hypothesis that fear of being "suckered" underlays defection, not desire to capture the "free rider's" payoff (*greed*, in Coombs's terms). This hypothesis predicts that a money-back guarantee, which offers the individual protection against losing a contribution, will work better than enforcing a contribution. This means that opportunity for gain through defecting is removed. The failure of the money-back guarantee to increase contributing beyond the level of the standard dilemma suggests that Brubaker is wrong. Fear of loss through contributing is not the critical motivation underlying defection.

On the other hand, the relative success of the enforced contribution is consistent with the hypothesis that desire for gain through defecting is the motivation underlying defection. By hypothesis, enforcing a contribution if a public good is provided works to promote contributing by convincing people that the good will be provided and by removing the opportunity for free riding if it is provided. Free from the fear that the good will not be provided and that their contribution will be lost, people contribute under this rule because they see nothing to be gained by not contributing and, presumably, because they recognize the personal benefits available from the public good. Enforcing contribution should a public good be provided is, by this hypothesis, an institutional modification that is appropriately attuned to widespread, perhaps characteristic, human motivations.

Notes

This research was supported in part by the National Science Foundation under grants SOC-790651 to Orbell and Dawes and SES-8107818 to Simmons. The opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation. Further support was received from the Institute for the Study of Policy Choices, Utah State University, Logan, UT.

1. In our experiments with a dominant incentive to defect of $5, about 30% of the subjects characteristically cooperate. A pessimist, of course, might see such cooperation rates as high.

2. It was, we maintain, a public good for the state, the country, and the world at large.

3. Thus, these step-level games are not strictly social dilemmas, which require a dominant incentive to defect. Note that this absence of a dominant incentive does not preclude their classification in standard terms as "public goods" games. See Frohlich and Oppenheimer (1970) and van de Kragt, Orbell, and Dawes (1985).

4. Of course, there is the possibility of an endless regress here: if an individual is tempted to free ride due to believing the guarantee will work, he or she may infer that the others will also be tempted, which means that the contribution rate will be lower, which makes the guarantee more important, and so on.

5. The distribution of sessions over locations was as follows:

	Utah	Oregon
Standard dilemma	2	8
Money-back guarantee	3	4
Enforced contribution	5	5

6. The test can be found in Scheffé (1960). The computations were executed through the SPSSX one-way module and its Scheffé test option (SPSSX, 1983, p. 458).

7. The distribution of sessions over locations was as follows:

	Utah	Oregon
Standard dilemma	7	3
Money-back guarantee	4	3
Enforced contribution	5	5

8. We wish to acknowledge our debt to Amnon Rapoport for developing this approach and our pleasure that he found these step-level games of sufficient interest to develop his interpretation of the results. Palfrey and Rosenthal (1984) have developed an alternative interpretation based on mixed strategies. While their approach contains some elements identical to Rapoport's (e.g., solving for indifference between contributing and not contributing), the predictions following from it allow so many possible pure and mixed equilibria that we do not elaborate on it here.

9. The typographical error of omitting zero from this statement apparently had no effect; almost all subjects gave estimates for the three possibilities adding to one.

10. That the estimates of being deficient, critical, and redundant are not equal is a trivial finding ($F_{2,102} = 8.60$, $p < .01$). What we report as effects are all, technically, *interactions* with these three estimates.

11. See Goodman (1964).

12. In fact, this analysis is confounded with the earlier analysis of variance of repeated measures, which treated choice as the independent variable and estimates as the dependent one, rather than vice versa. We present both analyses, complete with significance levels, not because we believe in redundant analyses as a general desideratum—quite the opposite—but because each alone is flawed. In this context, we know of no overall best analysis. Certainly, an overall goodness-of-fit statistic, given directional support but striking discrepancy from predictions, would be inappropriate.

References

Axelrod, Robert, and William D. Hamilton. 1981. The Evolution of Cooperation. *Science*, 211: 1390–96.

Brubaker, Earl R. 1975. Free Ride, Free Revelation or Golden Rule? *The Journal of Law and Economics*, 18:147–61.

Caldwell, Michael D. 1976. Communication and Sex Effects in a Five-Person Prisoner's Dilemma Game. *Journal of Personality and Social Psychology*, 33:273–80.

Coombs, Clyde. 1973. A Reparameterization of the Prisoner's Dilemma Game. *Behavioral Science*, 18:424–28.

Dawes, Robyn M. 1980. Social Dilemmas. *Annual Review of Psychology*, 31:169–93.

Dawes, Robyn M., Jeannie McTavish, and Hariett Shaklee. 1977. Behavior, Communication and Assumptions about Other Peoples' Behavior in a Commons Dilemma Situation. *Journal of Personality and Social Psychology*, 35:1–11.

Edney, Julian, and Christopher S. Harper. 1978. The Commons Dilemma: A Review. *Environmental Management*, 2:419–507.

Frohlich, Norman, and Joe Oppenheimer. 1970. I

Get by With a Little Help From My Friends. *World Politics,* 23:104–20.
Goodman, Leo A. 1964. Simple Methods for Analyzing Three-Factor Interactions in Contingency Tables. *Journal of the American Statistical Association,* 59:319–52.
Jerdee, Thomas H., and Benson Rosen. 1974. Effects of Opportunity to Communicate and Visibility of Individual Decisions on Behavior in the Common Interest. *Journal of Applied Psychology,* 59:712–16.
Olson, Mancur. 1965. *The Logic of Collective Action.* Cambridge: Harvard University Press.
Orbell, John M., Peregrine Schwartz-Shea, and Randy T. Simmons. 1984. Do Cooperators Exit More Readily than Defectors? *American Political Science Review,* 78:147–62.
Palfrey, Thomas R., and Howard Rosenthal. 1984. Participation and the Provision of Discrete Public Goods: A Strategic Analysis. *Journal of Public Economics,* 24:171–93.
Rapoport, Amnon. 1985. Provision of Public Goods and the MCS Experimental Paradigm. *American Political Science Review,* 79:148–55.
Rapoport, Anatol. 1974. The Prisoner's Dilemma: Recollections and Observations. In Anatol Rapoport, ed., *Game Theory as a Theory of Conflict Resolution.* Boston: D. Reidel Publishing Co.
Scheffé, Henry A. 1960. *The Analysis of Variance.* New York: John Wiley.
Simmons, Randy T. 1980. *The Logic for Cooperation? An Investigation of the Free Rider Hypothesis.* Ph.D. diss., Department of Political Science, The University of Oregon.
Smith, Adam. 1976. *The Wealth of Nations.* Chicago: The University of Chicago Press. (Original work published in 1776).
SPSSX. 1983. *SPSSX Users' Guide.* New York: McGraw-Hill.
van de Kragt, Alphons J. C., John M. Orbell, and Robyn M. Dawes. 1983. The Minimal Contributing Set as a Solution to Public Goods Problems. *American Political Science Review,* 77:112–22.
van de Kragt, Alphons J. C., John M. Orbell, and Robyn M. Dawes. 1985. Reply to Calvert and Wilson. *American Political Science Review,* 79: 823–24.
van de Kragt, Alphons J. C., John M. Orbell, and Robyn M. Dawes, with Sandy R. Braver and L. A. Wilson II. 1986. Doing Well and Doing Good as Solutions to Public Goods Problems. In Henke Wilke, David Messick, and Crystel Rutte, eds., *Experimental Social Dilemmas.* Frankfurt am Main: Lang GmbH.

Robyn M. Dawes is Head of the Department of Social and Decision Sciences and Professor of Psychology, Carnegie-Mellon University, Pittsburgh, PA 15213.

John M. Orbell is Professor of Political Science, University of Oregon, Eugene, OR 97401 and Visiting Professor of Political Science, Utah State University, Logan, UT 84322.

Randy T. Simmons is Associate Professor of Political Science, Utah State University, Logan, UT 84322.

Alphons J. C. van de Kragt holds a Post Doctoral Fellowship at and is an Instructor in the Department of Social and Decision Sciences, Carnegie-Mellon University, Pittsburgh, PA 15213.