GROUP SIZE EFFECTS IN PUBLIC GOODS PROVISION: THE VOLUNTARY CONTRIBUTIONS MECHANISM*

R. MARK ISAAC AND JAMES M. WALKER

This paper examines the relationship between variations in group size and "free-riding" behavior in the voluntary provision of public goods. We examine experimentally two pertinent concepts: the marginal return to an individual from contributions to the public good, and the actual number of members in the group. Our results strongly support a hypothesis that increasing group size leads to a reduction in allocative efficiency when accompanied by a decrease in marginal return from the public good (as from crowding or an association of large groups with imperceptibility of marginal benefits). Our results do not support a pure numbers-in-the-group effect.

I. INTRODUCTION

Much of the literature regarding the provision of public goods focuses on the problems associated with underrevelation of demand and the relationship of such "free-riding" behavior to variations in group size. This paper offers experimental evidence of individual and group behavior obtained from laboratory experiments designed to focus on this issue. We operationalize and test two separable concepts that, standing alone or in conjunction with each other, are integral factors in any attempt to analyze the relationship between group size and public goods provision: (1) the marginal return to an individual from contributions to the public good; and (2) the actual number of participants in the group.

In Section II we give an overview of the decision faced by participants. This is followed by a demonstration of the relationship of the concepts listed above to variations in group size, both in the voluntary contributions mechanisms and also in relation to similar concepts in n-person prisoners' dilemmas. Section II concludes with a summary of the predictions of formal theoretical models that are pertinent to the specific environment we investigate. The experimental design and a discussion of instructions reviewed by participants are presented in Section III. In Section IV we present our results, providing observations on both individual and group effects. In summary, we argue that, at least for the range of variables we have investigated, the standard argument that

*The financial support of the National Science Foundation (grant #SES8310096) is gratefully acknowledged. We also wish to thank Eric Maskin and two anonymous referees for comments on an earlier draft.

"large" groups have a more difficult time providing public goods than "small" groups appears to be strongly supported when that distinction in group size is driven by reductions in the marginal per capita return to an individual from contributions to the public good. We find no obvious separate effect from the actual number of participants. Section V offers a discussion of the relationship of our results with other research. Section VI concludes the paper with some thoughts for further experimental and theoretical research.

II. Group Size Concepts

A. The Choice Environment

The experiments reported here employ a computerized version of the voluntary contribution public goods provision mechanism. More detail on the experimental design and instruction is provided in the next section, but the gist of the process is as follows.[1] Each individual in a group of size n faces ten "investment" decision periods. In each period the experimenters endow each participant (i) with Z_i tokens. Each token will be invested in an "individual exchange" (where it pays the individual \$0.01 with certainty), or in a "group exchange." Let m_i represent individual i's contribution of tokens to the group exchange in a given period. The group exchange is the public good in that each individual receives a payment of $(1|n)G(m_i + \Sigma\, m_j)$ cents, where $G(\cdot)$ is an appropriately specified function and $\Sigma\, m_j$ represents the sum of the contributions of everyone else except person i. In fact, the $G(\cdot)$ function was chosen so that the Pareto optimum (in this experiment defined simply as the outcome that provides the greatest total monetary payout from the experimenters to the subjects) was for every individual always to invest all tokens in the group exchange (i.e., to set $m_i = Z_i$). Note that an individual is free to divide his tokens between the two types of exchanges in any one period, but is not allowed to carry over tokens from one period to the next.

Because contributing tokens to either of the exchanges results in a well-defined monetary payoff, we can write a representative individual's utility function in any one period as

$$U_i(Z_i - m_i + (1|n)G(m_i + \Sigma\, m_j)).$$

If the individual views each trial as a single decision problem, this

1. Copies of the complete instructions are available from the authors upon request.

would be represented as

$$\max_{0 \leq m_i \leq Z_i} U_i(Z_i - m_i + (1|n)G(m_i + \Sigma\, m_j)).$$

B. Interpreting Changes in Group Size

For many years researchers in the area of voluntary public goods provision have discussed separating the components that make up larger versus smaller groups for purposes of providing public goods. Some examples are Olson [1965], Buchanan [1968], Frohlich and Oppenheimer [1970], and Chamberlin [1974]. Investigating the effects of changing group size in the context of the voluntary contribution mechanism is, in some ways, conceptually parallel to the issue faced by analysts of n-person prisoners' dilemma (NPD) games. NPD researchers who have addressed this issue include Hamburger et al. [1975], Bonacich et al. [1976], Komorita [1976], Dawes [1980], Hardin [1982], and van de Kragt [1984]. Our intention here is to demonstrate three possible ways one might interpret the concept of changes in group size within the voluntary contributions mechanism, and then to investigate the similarity of these concepts to those found in the NPD literature. While the prisoners' dilemma and the voluntary-contribution public goods mechanism can easily be described with similar normal forms, the two institutions vary considerably in their extensive form.[2] The experimental literature focusing on group size effects is much more extensive in NPD than in public goods research.

In many instances, discussion of public goods provision has distinguished between "pure" versus "impure" public goods. Our interpretation of group size is partially motivated by this distinction. A pure public good can be characterized as having perfect nonrivalry in consumption. Thus, increasing group size does *not* reduce the marginal benefit of the public good to other consumers. Counter to this type of public good is that of the "impure" public good which can be jointly consumed but in which increases in group size tend to diminish the marginal benefit to all consumers (that is, there is a crowding effect). Given this latter interpretation, there is an interaction between the number of persons involved in the voluntary contributions process and the benefit received from the group good by any one individual (in our environment, $(1|n)G(\cdot)$). Versions of this interaction have been noted by, among others, Olson, Frohlich and Oppenheimer, Hardin, and van de Kragt.

2. The normal forms are not identical because the voluntary contributions mechanism has a strategy space larger than the two-element "defect, cooperate" pair typical of prisoners' dilemmas.

To continue to take this interpretation of group size into our laboratory environment, note that there are four relevant parameters: (i) marginal benefits from the individual exchange, (ii) the functional form of $G(\cdot)$, (iii) the number of persons participating (n), and (iv) the endowment of tokens to each individual (Z_i). To see how our mechanism admits of three different interpretations of a group size effect, consider our base parameterizations: marginal benefits from the individual exchange equal to one cent; $G(\cdot) = 3(\Sigma_i m_i)$; $n = 4$; and $Z_i = 25$. If this were considered a "small group" experiment, how would one conduct a "large group" experiment? One interpretation would simply be to hold $G(\cdot)$ constant and increase n. As an example, consider raising n to ten. Such an alteration combines two separable effects. First, the number of individuals in each decision group increases from four to ten. Second, $(1|n)G'(\cdot)$, the *marginal* per capita return (MPCR) of a contribution to the public good, *falls* from 0.75 to 0.3. Alternatively, as one increases n from four to ten, the specification of $G(\cdot)$ could be changed to $G(\cdot) = 7.5(\Sigma_i m_i)$. This second approach would allow n to increase, while holding MPCR constant at 0.75. Thus, we have identified three separate possible treatments relating to the way in which a group may be said to be large or small:[3]

1. The effect of altering MPCR but holding n constant;
2. The effect of altering n but adjusting the $G(\cdot)$ function so that MPCR remains constant;
3. The combination effect in which altering n also alters MPCR.

It is important to note that changes in any one parameter or combination of parameters within our mechanism have specific consequences for the decision environment. (The same is true of NPD games.) For example, altering n or MPCR will also change such values as total gains from unanimous group cooperation, total payoff from complete free riding, and the opportunity cost of contributing to the group good. For this reason, some up-front design choices must be made in order to conduct the experiments. As a rule, in the experiments reported here we (a) hold constant at one cent the payoff for contributing an additional token to the individual exchange, and (b) vary individual endowments (Z_i) with changes in n or MPCR so as to keep the per capita monetary value

3. While working on this paper, we became aware of the independent work of van de Kragt [1984] who addressed the same question (about what makes a group larger or smaller) from the perspective of the literature on prisoners' dilemmas.

of the Pareto frontier constant across all experiments.[4] The impact of this particular choice of experimental constants is a behavioral question that is an obvious candidate for future research. The exact nature of the payoff structure for each of our experimental designs is discussed in greater detail in Section III.

C. Comparison of Group Size Concepts for NPD Games

As the voluntary contributions mechanism and n-person prisoner's dilemma are similar in normal form, theoretical and experimental NPD research has had to grapple with a similar problem of deciding what makes a prisoner's dilemma large or small. The concept of the actual number of participants, n, is directly parallel with our discussion above. The concept of MPCR is not. This is because in NPD games participants typically have only two choices in their strategy set: cooperate or defect. There is no identical marginal effect of moving a single token from the individual to the group exchange. Nevertheless, NPD games do have analogous concepts of incremental incentives and, like MPCR, they are intertwined with concepts of group size.

A recent paper by van de Kragt [1984] is devoted to examining concepts of group size in NPD games, extending analyses by other NPD researchers. In van de Kragt's terminology, call $g(C(n))$ the gain each group member receives if all n choose to cooperate. He shows that as n increases either (i) $g(C(n))$ also increases, or (ii) if $g(C(n))$ is held constant, some other parameter must decline. One of the values that could decline in order to hold $g(C(n))$ constant as n increases is the "externality" of the dilemma (which is the change in any individual's payoff as any one person decides to cooperate or defect). This latter possibility is interpretable as the discrete equivalent in a prisoners' dilemma of MPCR declining as n increases in the voluntary contributions mechanism.

D. Hypotheses of Group Size Effects for Our Environment

The previous section developed three possible interpretations for our environment of what it means to make a group larger or smaller. These three interpretations are composed of what we call the marginal per capita return component (MPCR) and the pure numbers in the group component (n). The textbook conjecture is

4. One alternative would have been to hold constant the number of tokens given to each individual. This would have standardized wealth at the outcome where every individual contributes all tokens to the private good.

that "large" groups have a harder time providing the optimal level of the public good than do "small" groups, where now "large" and "small" can be more precisely defined for our experiments. A logical next step is to ask whether there exist any other received theories that would generate predictions for our economic environment regarding changes in group size.

First, there is the literature of complete information noncooperative game theory. Consider as maintained assumptions two facts that were true about *all* of our chosen parameterizations: (i) the return from a token invested in the individual exchange is always one cent; and (ii) the MPCR from contributing to the group exchange is always strictly less than one cent. Then, this decision process generates several predictions from standard noncooperative game theory. For example, interpreted as a single-period decision process, there is a dominant strategy for each individual to contribute zero tokens to the public good (the group exchange). Further, in the finite repeated-game framework [Friedman, 1977] zero contribution by each individual is the unique multi-period Nash equilibrium. Thus, standard noncooperative game theory predicts identical outcomes for any combinations of MPCR and n satisfying the restrictions stated above. That is, there should be *no* observable group size effects when changes are limited to this range.

There are also alternatives to the standard noncooperative model. In particular, Kreps et al. [1982] have formulated a model that predicts cooperation in early periods of a *finitely* repeated prisoners' dilemma. In the standard model, complete noncooperation is predicted as a Nash Equilibrium even in early periods through a backwards induction argument. In the Kreps et al. model, the backwards induction argument does not hold because the players begin the experiment in a state of incomplete information on rivals' payoff matrix or rationality. Indeed, results from several public goods experimental series have demonstrated a similar pattern of early-period cooperation with eventual decay toward the standard free-riding outcome [Isaac, McCue, and Plott, 1985; Kim and M. Walker, 1984; Isaac, J. Walker, and Thomas, 1984—hereafter IWT].[5] Fudenberg and Maskin [1986] describe conditions under which a sufficiently long finitely repeated prisoner's dilemma can (under incomplete information) have average payoffs that approximate any (one shot) individually rational payoff vector. Unfortunately, the incomplete information models to this point do

5. Not all of the experimental environments match that specified in Kreps et al. Specifically, only IWT had a known, finite end-period.

not provide operational quantitative predictions for our environment beyond the qualitative prediction of positive early-period contributions collapsing to zero contributions in the final period.[6]

If any of the standard received theories should turn out not to be good predictors of the *level* of contributions in our voluntary contributions environment, then one might expect to find (also at variance with the standard theories) treatment effects from either MPCR or n. The purpose of this paper is to examine whether either the marginal per capita return from contributing to the public good or the number of persons in the decision group can be shown to have a significant and consistent impact on the levels of individual and group contributions in a voluntary contributions environment.

III. Experimental Design and Parameters

A. Subjects and Experimental Setting

The experiments reported in this study were conducted using subjects drawn from a population of undergraduate students at the University of Arizona. The students were currently enrolled in lower-level economics courses. All of the experiments were conducted using *experienced* subjects. These subjects were randomly chosen from a pool of those who had participated in previous experiments employing the same decision process. All subjects were volunteers who (in prior recruiting) had received a brief explanation of what it meant to be in an economics experiment.[7] Emphasis had been placed on the fact that no special background in economics or the use of computers was needed to participate. In recruiting experienced subjects, we never replicated a particular group from previous experiments.

The experiments were conducted using the PLATO computer system. This system allows for minimal experimenter-subject interaction during experimental sessions and also insures that all subjects see identical instructions and examples for a given experimen-

6. The existence of a group size effect consistent with the incomplete information models appears to be a possibility. For example, suppose I believe that there is a small (independent, binomial) probability that any one of my fellow players "enjoys cooperating." Then, the chances that I face at least one such rival changes with the number of rivals. Again, the differences in normal and extensive forms makes the parallels less than exact and, to our knowledge, this literature has been developed primarily in the context of two-person games.

7. We have observed that the proportion of participants who wish to revolunteer is over 95 percent. Also, more than 90 percent of the participants in the experiments reported in this paper had been exposed to both the low and high MPCR payoff conditions in prior experiments. Of the small number who had seen only one MPCR condition, they were well mixed between low and high MPCR.

tal design. The use of the computer also facilitates the accounting process that occurs in each decision period and minimizes subjects' transactions costs in making decisions and recalling information from previous decisions.

B. The Decision Mechanism: A Summary

At the beginning of each experimental session, participants were told that they would be participating in an economic market in which they would make investment decisions. It was explained to the participants that the computer system was used only to instruct them in the specifics of the decision problem they would face and act as the means for transacting their decisions. It was emphasized to each participant that there would be no communication with other participants during the experiment except for communications allowed through the computer. It was also emphasized to participants that their individual investment decisions would not be known by other participants in the market. The subjects faced individual computer terminals with side-board blinders used to gain as much privacy and anonymity as possible.

The programmed instructions described to the participants the following decision problem: given a specific endowment of tokens (resources) participants faced the decision of allocating their tokens between an individual exchange (private good) and a group exchange (public good).[8] The individual exchange was described as an investment that paid to the investor one cent for each token invested. The group exchange was explained to the participants as an investment that yielded a specific return per token to the individual (as well as yielding the same return to all other participants). Thus, the payoff a participant received from the group exchange was explained to depend upon his own investment in the group exchange as well as upon the investment in the group exchange by all other participants. It is important to note that, although the total payoff from the group exchange is dependent upon the investment of others, the MPCR does not vary. From any one contributors' point of view, there is no necessity to conjecture about the marginal benefits from incremental changes in contributions to the public good. The payoff from the group exchange was reported to each participant in the form of a table that gave group

8. The phrases "individual exchange" and "group exchange" were the ones used in the experiment. The reference to private and public goods is provided here to be more consistent with standard terminology.

and individual returns from the group exchange for various investment levels (from zero up to the total tokens owned by the group).

The information position of each participant can be described as follows. First, each participant knew his own endowment of tokens for each decision trial and the total number of tokens for the group. He did not know the specific allocation of tokens to other participants.[9] Second, participants knew the exact size of the group and that each participant's return from the group exchange was identical. Each participant knew with certainty his own return from the private exchange. Third, each participant knew that there would be ten decision trials and that his endowment for each trial would be equal. Finally, it was explained that the monetary gains from each trial were binding and total payments to the participant equaled the sum of his return from the group and individual exchanges totaled over all ten trials. At the end of each trial the participant received information on his return from the individual and group exchange. Each was also told the total number of tokens contributed by the group to the group exchange. Before making an investment decision in any one trial, a participant could obtain this same information for all previous trials.

The participants faced two consecutive series of ten decision trials. This provided (to our knowledge) the first *intra*group comparison of the effects of the marginal per capita return on provision of the public good. In one series, this parameter was set equal to 0.3; in the other, it was set equal to 0.75. (These two values were chosen to correspond to the high and low MPCR conditions in previous research reported in IWT). As a check for any possible sequencing effects, the order was (0.3, 0.75) in half of the twelve two-series experiments and (0.75, 0.3) in the other half. The multiple replication of design allows for a check on the robustness of the *inter*group MPCR effects encountered in previous work.

Because participants are aware that each series consists of precisely ten trials, we have not only a ten-period finitely repeated process but also, from period 10, a true "single-period" decision context in which a substantial opportunity for "learning" has already occurred. In the experiments reported here, participants in the first series of ten trials did not know with certainty of the existence of a second series until after the first series completion. However, they were experienced participants and had presumably observed that in their previous participation the maximum time for

9. In ongoing research, we are investigating the effects of altering this information condition.

which they had been recruited (two hours) allowed for two ten-period sequences. Therefore, to be conservative, we shall analyze as a single-period decision opportunity *only* the data from the ten-period of the second (last) series.

C. Experimental Parameters

The results reported here are based on twenty-four experiments. Each of twelve experimental sessions consisted of two ten-period experiments, alternating MPCR between the two sequences. There were twelve four-person experiments and twelve ten-person experiments. As a result, and based upon the earlier exposition of concepts of group size, this design can be represented as a four-cell process among the combinations of MPCR and n. These combinations are represented in Table I, with cell $4L$ indicating four-person experiments using the lower MPCR value (0.3), and so forth. In presenting the results in the next section, we shall focus on the separate simple effects of MPCR and n, and then discuss those results in the context of our three interpretations of group size.

IV. Empirical Results

We turn now to a discussion of the experimental results. We organize the presentation around a series of summary observations. In certain cases, the reader is referred through the footnotes to the results of statistical tests regarding the results, but these tests are not presented in detail here. We present observations on both individual behavior and on the aggregate provision of the public good. In terms of observations related to individual behavior, we shall focus upon the number of individuals who are what we have (somewhat arbitrarily) defined as "strong free riders." An individ-

TABLE I
Experiment Parameters

Experiment type	Group size	Group payoff function	MPCR	Individual tokens per period (Z_i)	Number of experiments
$4L$	4	$1.2(\Sigma m_i)$¢	0.30	62	6
$4H$	4	$3.0(\Sigma m_i)$¢	0.75	25	6
$10L$	10	$3.0(\Sigma m_i)$¢	0.30	25	6
$10H$	10	$7.5(\Sigma m_i)$¢	0.75	10	6

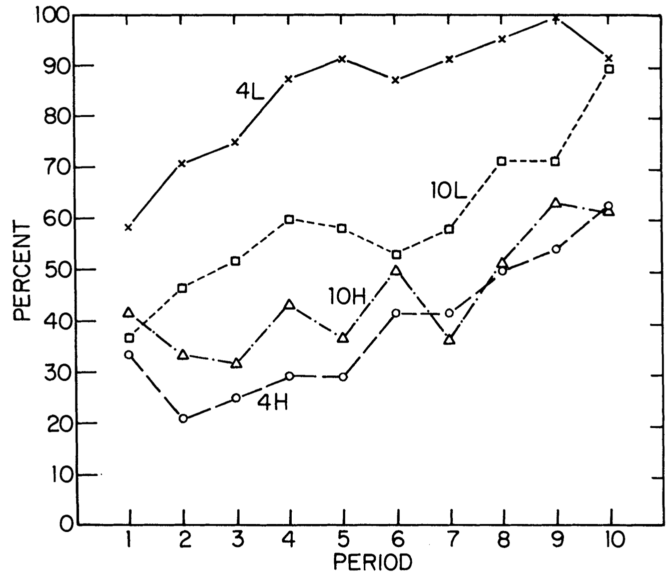

FIGURE I
Mean Percent of Individuals Acting as Strong Free Riders

ual is said to be a strong free rider if his contribution to the public good is less than one third of his tokens (one third of the Pareto optimal contribution). At the suggestion of readers of earlier versions of this work, we checked the robustness of these findings against an alternative definition of strong free rider which requires that we count only those persons who contribute precisely zero tokens. None of the qualitative conclusions we are to report reverse using this alternative specification. Finally, we should comment upon evidence regarding possible sequencing effects due to our intragroup sequencing of MPCR. We found no indication of an effect of sequence upon our reported observations, so we do not disaggregate our presentation by sequence.[10]

The main conclusions of our research can be drawn from the data presented in Figures I and II and Table II. Figure I displays the period-by-period mean percentage of individuals acting as strong free riders for each of the four treatments (*4L*, *4H*, *10L*, *10H*). Figure II presents mean period-by-period contributions to the public good (as a percent of optimum) for the same configuration.

10. This is not to say that there were no sequencing effects at all, only that we could find none that altered the conclusions we report.

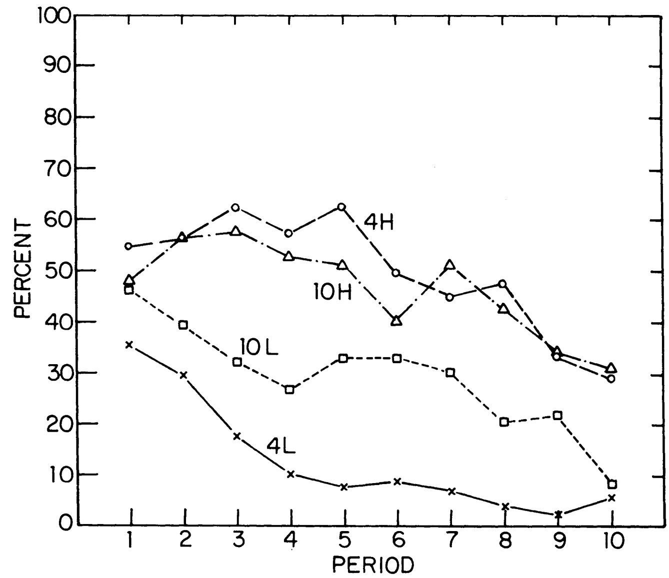

FIGURE II
Mean Percent of Tokens Contributed to the Public Good

Table II contains the data from the tenth period of each second-sequence experiments, which comprise our sample of "end-period" observations.

OBSERVATION 1. Lowering the MPCR from 0.75 to 0.3 appears to increase significantly the incidence of free-riding behavior.

In periods 2 through 10, the mean percent of strong free riders is greater in *both* groups with a low MPCR than for *either* group with a high MPCR. The same is true for *all* periods for mean contribution to the public good as a percent of optimum. Notice further that the low MPCR experiments display more clearly a "decay" phenomenon across iterated decisions.

In the true end-periods, the MPCR effect is still present. On average, low MPCR groups contributed only 3.65 percent of their tokens (with an average of 95 percent acting as strong free riders) while groups with the higher MPCR averaged 26.35 percent contribution (with an average of 68.3 percent acting as strong free riders).

Our sequencing design also allows for an intragroup comparison of the effects of MPCR. For each period, there are twelve paired intragroup data comparisons. Figures IIIa and IIIb illustrate how

TABLE II
END-PERIOD RESULTS

Experiment type	Replication	Percent of tokens contributed to the public good	Number of strong free-riders	Number of persons contributing zero
4*L*	1	0.0%	4	4
	2	0.0%	4	4
	3	0.0%	4	4
4*H*	1	25.0%	2	2
	2	52.0%	2	1
	3	10.0%	4	2
10*L*	1	0.4%	10	9
	2	11.6%	8	8
	3	10.0%	9	6
10*H*	1	25.0%	7	6
	2	21.0%	7	7
	3	25.0%	7	6

many of the twelve comparisons for each period resulted in the 0.3 MPCR having more strong free riders (Figure IIIa) or lower percent contributions to the public good (Figure IIIb).

OBSERVATION 2. There are weak, if any, effects of changes in group size from four to ten (holding MPCR constant).

From both Figures I and II, an interesting pattern emerges. For groups with MPCR equal to 0.75, there is no clear effect due to n. The 4*H* and 10*H* graphs are very close and, in fact, cross each other repeatedly. On the other hand, there is a consistent tendency for n to make a difference in groups with the low MPCR. In ten of ten periods, the mean data suggest more free riding in the *smaller* group ($n = 4$) than in the larger group ($n = 10$) using the criterion of either group or individual behavior.

In the end-periods group size appears to have a negligible effect. The four-person and ten-person groups contributed an average of 14.5 percent and 15.5 percent of their tokens to the group exchange, respectively. Their proportions of strong free riders were 83.35 percent and 80 percent, respectively.

V. INTERPRETATION OF RESULTS

The results just presented lead to some interpretations that may be used in evaluating several formal and informal hypotheses about public goods provision. Also, the results can be checked for consistency with conclusions drawn by other researchers.

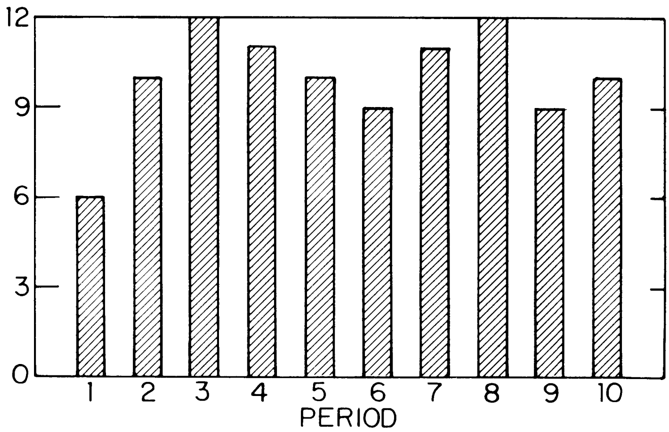

FIGURE IIIa
Period-by-Period Paired Intragroup Comparisons: Low MPCR Has More Strong Free Riders

A. Interpretation of Group Size Effects in Public Goods Research

To what extent do the data support or dispute the hypothesis that "large" groups have a more difficult time providing public goods than "small" groups? Without ambiguity, groups with a lower MPCR had less success in providing optimal levels of the public good. Thus, to the extent that increasing group size decreases

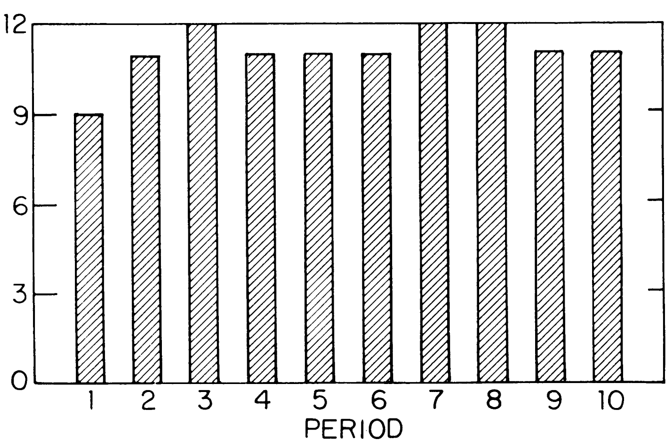

FIGURE IIIb
Period-by-Period Paired Intragroup Comparisons: Low MPCR Has Lower Percent of Tokens Contributed to Public Good

MPCR (through crowding, for example), these results are consistent with the traditional view. On the other hand, controlling for MPCR, we found that changing n had a weak and ambiguous effect. In fact, to the extent that there was any difference in the data, it was in the direction of the large group being relatively more successful. Therefore, our data do not suggest that more group decision makers *in and of itself* makes a group less likely to obtain efficient levels of the public good.[11]

As mentioned earlier, one obvious way for a group to go from small to large is for both n and MPCR to change (more people increases n and produces crowding). The results described above would suggest that public goods provision would be closer to efficient in the small group. In our four-treatment design, we have a direct test of this combined effect by comparing just two cells: $4H$ with $10L$. Figure IV displays the comparison of the mean pattern of total group contributions. The combined effect of the two components of group size is clearly as expected: the groups with n equal to four and with the large MPCR consistently contributed more, on average, to the public good.

To our knowledge, the only other experimental research of the voluntary contributions mechanism to investigate group size is reported by Marwell and Ames [1979]. There are many differences between their experimental design and ours. Among the differences are the following: their experiments were "one shot," conducted by telephone, and used a "provision point" payoff function.[12] In addition, all subjects *actually* participated in groups of four persons. However, the participants were sometimes told that they were participating in groups of 80 persons. In one treatment, Marwell and Ames altered the reported group size and, in the process, lowered a parameter in their design that is analogous to what we have called MPCR.[13] Thus, this change is analogous to our treatment where both n and MPCR are altered. In summary, Marwell and Ames found that average contributions were lower in the larger groups, but the difference was not statistically significant.

11. This does not address the issue as to whether small or large groups are more likely to form. Recall that in these experiments the costs of organizing and assembling the group were borne by the experimenters.

12. A "provision point" payoff function is one with a discontinuity at some minimum required level of contribution. If total contributions fall short of the provision point, none (or substantially less) of the public good is provided.

13. In the Marwell and Ames experiments, the "group good" payoff function for the large groups required an investment twenty times as large as that of the small group for the same per capita return. Disregarding the complications presented by the use of the provision points, this change in the payoff function can be interpreted as a decline in MPCR.

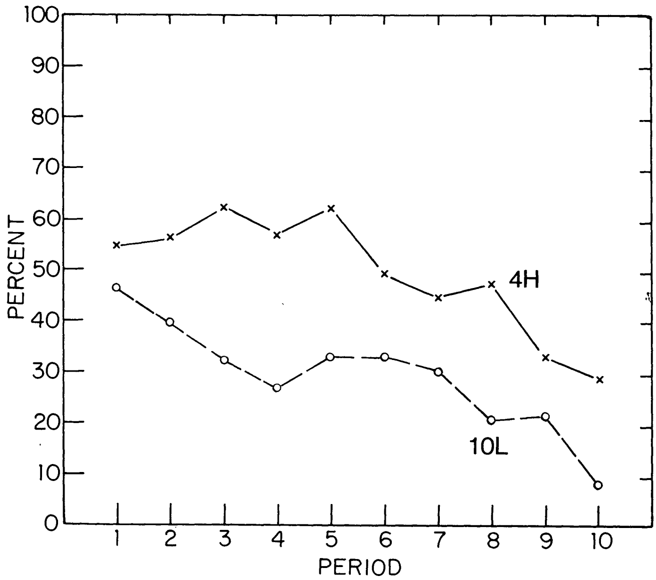

FIGURE IV
Mean Percent of Tokens Contributed to the Public Good: 4*H* vs. 10*L*
X = 4*H*; O = 10*L*

B. Game-Theoretic Economic Models

The *end-period* results reported above provide a test for the clear-cut predictions of the single-period dominant strategy model. The prediction is unambiguous that, for any of the four treatments, no participant should contribute any tokens to the group exchange. In fact, 59 out of 84 people contributed precisely zero, while 68 out of 84 contributed fewer than one third of their tokens. (Total contributions in the end-periods averaged approximately 15 percent of the optimal level.) While the proportion of persons making decisions consistent with the dominant strategy is large, what of the remainder? As noted above, there is a residual MPCR effect. Thirty-five of forty-two persons contributed zero in groups with the low MPCR, while twenty-four of forty-two contributed zero when faced with a high MPCR.[14] There appears to be no independent effect of n.[15]

14. The Z-statistic on difference in proportions is 2.625.

15. Seventeen of twenty-four persons in the four person groups contributed precisely zero (70.8 percent). Forty-two of sixty persons in the ten-person groups contributed precisely zero (70.0 percent). The Z-statistic on the difference in proportions is 0.072.

The results across all periods are not supportive of the multi-period Nash Equilibrium prediction of zero contribution in every period (based upon a backwards induction argument). Instead, the experiments uniformly begin with positive contributions (ranging from 1 to 90 percent of optimum, averaging 46 percent of optimum) followed by a tendency for contributions to decay.[16] This decay pattern is consistent with the experimental results cited by Kreps et al., and it suggests that the incomplete information models should also be a fruitful line of theoretical inquiry for public goods research.

C. Prisoners' Dilemma Research

Comparison of our results with those from NPD research is somewhat difficult because of the important differences in the two institutions and in other matters of experimental design. However, several studies have addressed questions analogous to ours, and the reader is referred to the footnotes for more complete treatment of the design differences. Hamburger, Guyer, and Fox [1975] compare groups of three and seven persons. Their payoff functions were nonlinear and hence had nonconstant MPCR. However, our interpretation is that the expected value of the MPCR was greater for the smaller group. Thus, their change from three to seven combines an increased n with a decreased (expected) MPCR. Consistent with our results, they find more cooperation with the small groups. Bonacich et al. [1976] conducted experiments separating the concepts of n (they looked at groups with three, six, and nine persons) and incentives to cooperate. Our interpretation is that there is no design series precisely parallel to ours, but their results appear to be analogous to what we have found. In one series, they control for the level of the externality (their version of MPCR), while allowing n to increase.[17] In general, they found that, when controlling for their set of incentives to cooperate, an increase in n led to *more* cooperation. Komorita et al. [1980] tested groups with both four and eight persons. For each group size the slope of the "externality" function (analogous to MPCR) was found to have a significant effect. No significant separate effect of pure group size was found. Thus, this

16. Decay is more pronounced with the low MPCR.

17. The Bonacich et al. paper is a very systematic investigation of changes in group size in NPD games. In what is referred to as Rule A, they allow the externality from cooperating (similar to the MPCR) to remain constant while changing n. However, in order to maintain what they call "gains from cooperation" and "temptations not to cooperate," they alter the entry level of externality. In our voluntary contributions mechanism, a similar role was played by our alteration of number of tokens.

body of research appears to be very consistent with our findings of significant effects due to MPCR, but ambiguous effects of n.

A second prominent feature of our results was the presence of early positive levels of contribution decaying across time. Again, comparisons with NPD research are inexact (in particular, many of the repeated NPD experiments do not have known end-periods). However, Kelly and Grzelak [1972], Goehring and Kahan [1976], and Komorita et al. [1980] all indicate some type of decay across time for experiments with incentive structures analogous to having low MPCRs. Further, it is interesting to note that Rapoport and Chammah [1965] report longer declines in two-person games when subjects are not informed before every trial of the matrix of the other players. Notice that there is a similarity between this condition of not announcing the matrix in two-person games and the information condition of our four-person and ten-person groups. In our experiments, subjects knew the *total* of other contributions, but not the incentives or contribution of any other individual (except in the case in which the sum showed that everyone else contributed the maximum possible tokens).

VI. Conclusions

The results we report here are among the most consistently replicable experimental results we have encountered. The conclusion that a higher MPCR leads to greater efficiency in public goods provision (even when zero contribution is the identical prediction of the standard Nash models) is robust across intragroup and intergroup comparisons. We believe that these results have important implications for further theoretical work in at least two areas.

First, these results give significant evidence for separating the relative predictive power of different components of the concept of group size in public goods provision. At least for the 150 percent increase in group size from four to ten, there appears to be no support for a pure numbers argument relating increases in group size to increase in free-riding behavior. In fact, to the extent that there is any movement in the data, it is in the opposite direction.

Nevertheless, the traditional argument that larger groups will have more problems in providing public goods is strongly consistent with our findings that a decline in MPCR leads to more free riding. When one of the effects of increasing n is also to reduce MPCR (which was our original idea of going from a small to a large group) the unfavorable MPCR effect appears to dominate. Thus, our results are supportive of group size conjectures which suggest that larger groups have more problems with free riding than small groups because of factors which can be interpreted as declining or small MPCR. Some examples would be crowding or those large group environments in which an individual's marginal effect on the group decision is perceived to be negligible.

Second, we noted that our experimental environment can be modeled in normal form as an n-person repeated game with incomplete information and a common knowledge end-point. Thus, our results may be of interest to theorists working with such models. Overall, our results support the conclusion that the traditional zero-contribution prediction is a very powerful predictor in later periods, especially with a low MPCR for the public good. However, the process toward these outcomes does not conform to the usual backwards induction model. First of all, contributions begin relatively high and tend to reach lower levels through gradual decay. Second, the level of MPCR significantly affects behavior, even though for the values we examined there is no difference in the single-period dominant strategy prediction or in the complete information Nash Equilibrium condition. Even in the true end-periods, we observe a nontrivial minority of individuals (primarily in the high MPCR condition) contributing significant amounts to the public good.

We have several conjectures as possible causes for the observed behavior:

1. The concept of MPCR, which involves incentives at the margin, also has implications for the total reward conditions. Table III describes three total reward (or "wealth") values for each of the design configurations. The second column ("W") is the per period

TABLE III
REWARD CONDITIONS FOR COOPERATION AND FREE RIDING

Experiment type	W	M^1	M^2
$4L$	\$0.75	\$0.62	\$1.18
$4H$	\$0.75	\$0.25	\$0.81
$10L$	\$0.75	\$0.25	\$0.92
$10H$	\$0.75	\$0.10	\$0.78

W - Per person reward if all of the group's tokens are invested in the group exchange.
M^1 - Per person reward if none of the group's tokens are invested in the group exchange.
M^2 - Reward to a single participant for contributing zero tokens when *all* other participants contribute all possible tokens.

monetary reward to a single individual at the Pareto frontier. As was discussed earlier, this was standardized at 75 cents across all treatments. The third column (M^1) reports the per capita payoff if all participants contribute zero tokens to the public good. The fourth column (M^2) reports the payoff to a single individual if he is the *only* person to contribute zero, while all others contribute all of their tokens to the public good. Note from these figures that the net *total* gain to an individual of the group moving from zero public goods provision to the optimal level is substantially larger with the higher MPCR, holding group size constant. The same is true for increases in n, holding MPCR constant. Regardless of where one chooses to standardize the wealth level, higher MPCR or greater number of persons in the group provide a greater *total* incentive for the group to attempt a tacitly cooperative mode of behavior.

2. The standard repeated game models assume complete information. This experimental design does not allow one individual to know with certainty the nature of anyone else's complete payoff structure because they are not told anyone else's payment from the individual exchange nor the exact distribution of tokens. The approach of several recent authors is to examine repeated games with incomplete information. Our decay phenomenon suggests that this approach is promising, although it cannot explain our results of positive contributions in the true end-period.

3. Our participants are not prompted in any formal maximization of a repeated game model. They are presented with the decision environment via a payoff chart for the returns for the group good. Perhaps even our experienced participants are still learning the Nash or dominant strategies.

4. Perhaps individuals retain a residual motive to altruism. This is consistent with positive contributions in the true end-period as well as with the fact that these positive contributions seem to be related to MPCR. That is, by the very nature of MPCR, altruism is relatively more "expensive" when MPCR is 0.3 than when it is 0.75.

We hope that relating these conjectures and our observations will prove useful to further development of theories of public good provision, as well as to the motivation of further experimental work. Empirical researchers may wish to check the robustness of our observations beyond our parameter space. Theorists may be able to incorporate insight from our experiments into game-theoretic models of multi-period decision processes. Both of these efforts should further our understanding of group size effects in public goods provision.

UNIVERSITY OF ARIZONA
INDIANA UNIVERSITY

REFERENCES

Bonacich, P., G. H. Shure, J. P. Kahan, and R. J. Meeker, "Cooperation and Group Size in the *N*-Person Prisoners' Dilemma," *Journal of Conflict Resolution,* XX (1976), 687–706.

Buchanan, James M., *The Demand and Supply of Public Goods* (Chicago, IL: Rand McNally and Company, 1968).

Chamberlin, John, "Provision of Public Goods as a Function of Group Size," *American Political Science Review,* LXVIII (1974), 707–16.

Dawes, Robyn M., "Social Dilemmas," *Annual Review of Psychology,* XXXI (1980), 169–93.

Friedman, James, *Oligopoly and the Theory of Games* (Amsterdam: North-Holland, 1977).

Frohlich, Norman, and Joel Oppenheimer, "I Get by with a Little Help from my Friends," *World Politics,* II (1970), 104–20.

Fudenberg, Drew, and Eric Maskin, "The Folk Theorem in Repeated Games with Discounting and with Incomplete Information," *Econometrica,* LIV (1986), 533–54.

Goehring, Dwight J., and James P. Kahan, "The Uniform *N*-Person Prisoners' Dilemma Game, Construction and Test of an Index of Cooperation," *Journal of Conflict Resolution,* XX (1976), 111–28.

Hamburger, Henry, Melvin Guyer, and John Fox, "Group Size and Cooperation," *Journal of Conflict Resolution,* XIX (1975), 503–31.

Hardin, Russell, *Collective Action* (Baltimore, MD: Johns Hopkins University Press, 1982).

Isaac, R. Mark, Kenneth McCue, and Charles R. Plott, "Public Goods Provision in an Experimental Environment," *Journal of Public Economics,* XXVI (1985), 51–74.

Isaac, R. Mark, James M. Walker, and Susan Thomas, "Divergent Evidence on Free Riding: An Experimental Examination of Some Possible Explanations," *Public Choice,* XLIII (1984), 113–49.

Kelley, Harold H., and Janusz Grzelak, "Conflict Between Individual and Common Interest in an *N*-Person Relationship," *Journal of Personality and Social Psychology,* XXI (1972), 190–97.

Kim, Oliver, and Mark Walker, "The Free Rider Problem: Experimental Evidence," *Public Choice,* XLIII (1984), 3–24.

Komorita, S. S., "A Model of the *N*-Person Dilemma Type Game," *Journal of Experimental Social Psychology,* XII (1976), 357–73.

——, J. Sweeney, and D. A. Kravitz, "Cooperative Choice in the *N*-Person Dilemma Situation," *Journal of Personality and Social Psychology,* XXXVIII (1980), 504–16.

Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson, "Rational Cooperation in the Finitely-Repeated Prisoners' Dilemma," *Journal of Economic Theory,* XXVII (1982), 245–52.

Marwell, Gerald, and Ruth E. Ames, "Experiments on the Provision of Public Goods. I. Resources, Interest, Group Size, and the Free Rider Problems," *American Journal of Sociology,* LXXXIV (1979), 1335–60.

Olson, Mancur, *The Logic of Collective Action* (Cambridge, MA: Harvard University Press, 1965).

Rapoport, Anatol, and Albert M. Chammah, *Prisoner's Dilemma: A Study in Conflict and Cooperation* (Ann Arbor: University of Michigan Press, 1965).

van de Kragt, Alphons, "Experimental *N*-Person Dilemmas and Group Size: An Elaboration" presented at the annual meetings of the Public Choice Society, 1984.