Cooperation Under the Threat of Expulsion
in a Public Goods Experiment

by Matthias Cinyabuguma, Talbot Page, and Louis Putterman[*]

Abstract

In a public goods experiment with the opportunity to vote to expel members of a group, we found that contributions rose to nearly 100% of endowments with significantly higher efficiency compared with a no-expulsion baseline. Expulsions were strictly of the lowest contributors, and there was an exceptionally strong fall-off in contributions in the last period, when the expulsion threat was unavailable. Our findings support the intuition that the threat of expulsion or ostracism is one device that helps groups to provide public goods.

JEL Classification: C91, H41, D23
Keywords: Public goods, collective action, punishment.

**Cooperation Under the Threat of Expulsion
in a Public Goods Experiment**

1. Introduction

The problem of free-riding on others' provision of a public good is frequently cited as a reason to have the state supply such goods and collect contributions in the form of taxes. But the state-based solution is not always feasible, nor is it necessarily always the best among feasible options. Some public goods are too local (the cleanliness of the sidewalk shared by a few shopkeepers) or are of interest to too small a constituency (a local opera company) to attract state involvement. In other cases (public radio, organized religion), there are ideological or other reasons to separate the state from the desired collective action. And in still other cases (village woodlots and irrigation schemes), the monitoring and enforcement capabilities of the state are small relative to those of the local citizenry. Individual contributions to a wide range of collective endeavors are subject to potential free-riding. Yet they are also responsive to a variety of measures that allow collective action to succeed to varying degrees. This paper studies experimentally one such measure, the ability of groups to discipline members by threatening to expel free riders from the group.

The voluntary provision of public goods has been studied in the laboratory in the form of the voluntary contribution mechanism (VCM) or public goods (PG) game. Early experiments established certain stylized facts: contributions tend to average somewhat over 50% of endowments in the first period of play, but to decay to low levels with repetition (Davis and Holt, 1993). Pre-play face-to-face communication is a powerful tool for fostering cooperation (Isaac and Walker, 1988a; Sally, 1995).[1] Cooperation can also be fostered by permitting players to target punishments at one another following their initial contribution decisions (Fehr and Gächter, 2000a, 2002, Carpenter and Matthews (2002), Sefton *et al*., 2002), although earnings are increased little if at all in this case, due to the costliness of punishment (Bochet *et al*., 2002; Page *et al*., 2004).

A mechanism that fosters cooperation in many real organizations, but that has not been studied in the laboratory until recently, is the exercise of the power to determine

---

[1] Recent experiments have extended this finding to audio-visual and on-line communications. See Brosig *et al*., 2003; Frohlich and Oppenhe1imer, 1998; Bochet *et al*., 2002.

who is in and who out of the relevant group. Page *et al*. (2004) found that when subjects helped to determine group composition by ranking one another as prospective partners, more cooperative subjects were able to form groups of their own and to sustain cooperation, while players only interested in their own payoffs had incentives to act cooperatively so as to be grouped with cooperative players. In this experiment, both contributions and efficiency were significantly higher in treatments with regrouping than in those in fixed groups.[2]

In Page *et al*., groups were periodically formed *de novo*, so that one can think only loosely in terms of "getting into" or "being ejected from" a group. The present study puts into sharper focus one aspect of membership determination, the right to expel members. We model a large group that provides its members with a non-rival, non-excludable good and can permanently expel members on the basis of past performance.[3] Real-world applications include partnerships, teams, and associations from which members can be expelled for non-fulfillment of their obligations. Commonalities also exist with a broader range of collective action environments, such as ordinary workplaces, and clubs and organizations that provide both public goods and private benefits including sociability (see the discussion of tie-ins between public and private goods in Olson, 1965).[4] Although our experiment involves a formal sanctioning process, this might also be viewed as an approximate representation of the less formal social sanctions that come into play in some of these settings.

Our paper adds to the emerging literature on mechanisms to sanction free-riding. It has long been understood by economists that it can be advantageous to be able to

---

[2] A public goods experiment by Ehrhart and Keser (1999) also had groups form endogenously, but because their subjects were free to join any group unilaterally, cooperative subjects had no protection from free-riders and little cooperation could be sustained. By contrast, Gunnthorsdottir, Houser, McCabe and Ameden (2002) found that when the experimenters put initial high contributors together, cooperation was sustained more than in randomly formed groups.

[3] Note that the public good in our experiment is non-rival and non-excludable to the members of a group, but in the repeated game of 15 periods, the public good becomes excludable insofar as an individual can be expelled from the group.

[4] In such examples, incumbent members can be excluded outright, or they can be subjected to more subtle forms of ostracism, perhaps with similar effect. In the ordinary workplace, the right to fire a shirking employee plays a central role in efficiency wage models (Shapiro and Stiglitz, 1984) and is considered an essential feature of firms by Alchian and Demsetz (1972), among other writers. Giving managers the right to fire employees was a major element in programs of transition from state socialism (Djankow and Murrell, 2002). In most firms, incentives and expulsion rights are structured hierarchically, so that the analogy to a public goods game may not at first seem apparent. Yet it is well known that collective action problems—i.e., those of cooperating toward a common goal—are present even in hierarchical firms.

credibly threaten an action that is costly to oneself and that would be irrational for a strictly payoff-maximizing individual to carry out, *ex post* (Schelling, 1960). Frank (1987), Hirshleifer (1987) and other economists have explored the evolutionary fitness of "the emotions as guarantors of threats and promises." Fehr and Gächter (2000b, 2002) found that the inclination to retaliate against "unfair" actions was widespread and beneficial to group outcomes in their experiments, even when punishing was not consistent with the maximization of own pay-off (for example, in one-shot play, or in the final period of finitely repeated play).[5]

To test whether preferences for punishment are present and credible to others, our experiment was designed so that voting to expel members from a group engaged in a public goods game was always costly to those so voting as well as to those expelled. Because voting itself is costly, the threat of expulsion is not credible if all subjects are payoff maximizers and there is common knowledge of their preferences.[6] Like Fehr and Gächter, we find that this costly punishment was in fact brought to bear, and that the threat of it occurring quickly became credible enough to deter most free riding. Unlike the punishment mechanisms studied by Fehr and Gächter and others (including ourselves), however, the expulsion right unambiguously raised efficiency in our experiment, even when the earnings of the expelled players are given equal weight in assessing it. This may be explained in part by the fact that the punishment ended up being more drastic,[7] but it helped too that a majority-determined punishment was more exclusively targeted at free riders, and that fewer subjects needed to be punished, partly because their punishment became public knowledge. Another finding is that contributions fell dramatically in the last period, in which the expulsion threat was absent, strongly suggesting that they were driven mainly by threats.

---

[5] This result was confirmed by Carpenter and Matthews (2002), Sefton *et al*. (2002), Bochet *et al*. (2002) and Page *et al*. (2003).

[6] This differentiates our design from the theoretical model of an n-person repeated prisoners' dilemma in Hirshleifer and Rasmusen (1989) where, as in our experiment, there is a potential cost to members from reducing the size of the group. The key to their demonstration that a strategy that calls for the ostracism of anyone who fails to cooperate is an equilibrium is that there are no gains from cooperation in the final period, because defection is a dominant strategy. Although this is also the case in our experiment, we make voting to expel another a costly choice, whereas they assume that it carries no direct cost, an assumption critical to their result. See also their discussion of settings in which the threat of ostracism is used to enforce cooperative behavior.

[7] In this respect, our result depends upon a specific configuration of parameters, leaving it unclear just how far it can be generalized.

By providing more evidence of a propensity to punish, and of the credibility in practice of that propensity, we contribute not only to the literature on sanctioning but also to a broader literature which has begun to provide a better understanding of behaviors in collective action environments. That emerging literature, including Andreoni (1988), Offerman, Sonnemans and Schram (1996), Fischbacher, Gächter and Fehr (2001), Kurzban and Houser (2001), Ahn, Ostrom and Walker (2002), Gunnthordottir *et al.* (2002), and Casari and Plott (2003), makes allowances for the presence of more than one type of agent and for a Bayesian-type solution to finitely repeated games (Kreps *et al.*, 1982). Our findings are difficult to explain unless some agents care about something in addition to monetary payoffs and/or unless others' beliefs accommodate the existence of agents of this type.[8]

The rest of the paper proceeds as follows. In section 2, we describe the set-up of the experiment, its rationale, and behavioral expectations under the contrasting assumptions of (a) common knowledge that all are payoff maximizers, and (b) uncertainty about player type, with a positive probability that some players have a preference for punishing free riders. In section 3, we describe and analyze our experimental results with respect to contribution decisions and earnings (efficiency). In section 4, we further analyze our results, focusing on voting decisions and their consequences. We summarize our findings and provide concluding comments in section 5.

2. Experimental Design and Rationale

In a VCM or linear public goods game, $n > 2$ subjects are each provided with an endowment of money or tokens that they must allocate between a private account and a group account. Typically, as in our experiment, a number of discrete allocation choices are possible. In a homogeneous design, all members receive the same endowment and get the same fixed pay-out m per dollar invested in the group account. When $m < 1 < nm$, we have a free-rider problem: a player concerned only with his or her own payoff is

---

[8] Strictly speaking, it is sufficient that enough agents believe that the non-payoff maximizing type may exist, even if none in fact do. However, it is difficult to see how such beliefs could be sustained in a world in which all are payoff maximizers. Many of the papers cited, including Bochet *et al.* and Page *et al.*, provide evidence that the "nonstandard" types do exist and indeed are numerous.

better off if all contribute their full endowments than if each contributes less, but better off still if everyone except himself contributes. If all are payoff maximizers, their iterated dominant strategies are therefore to contribute nothing.

In both our baseline treatment and our expulsion treatment, players started out in a group of sixteen subjects and received ten experimental dollars (hereafter, E$10) in each of fifteen periods of play.[9] One experimental dollar converted to $.035 at the end of the session, which included two fifteen period repeated game treatments, and average earnings were about $25 for a 90 to 100 minute session (a $5 minimum was guaranteed but never actually binding). In a given period, a subject i's earnings in experimental dollars were given by

$$(10 - C_i) + (0.2)\Sigma_{\text{all } j}\, C_j \tag{1}$$

where C_i is the amount i assigned (contributed) to the group account, $m = 0.2$, and $\Sigma_{\text{all } j}C_j$ is the sum of contributions to the group account by all members, i included.

In the baseline treatment, group composition remained fixed. Following each contribution decision, subjects were shown on their monitors a list of one another's assignments to the group account in the current period and in the previous two periods, each individual's average assignment so far, and his or her earnings in the current period and so far in the treatment, with each subject identified only by a randomly assigned letter that remained fixed, and with subject information displayed in a random order.[10] The task of allocating funds between the two accounts was repeated fifteen times, with the terminal period being known from the start to all participants.[11]

---

[9] Initial group size in our experiment is unusually large. Isaac and Walker (1988b) conducted VCM experiments in which group size was either 4 or 10 subjects and the marginal per capita return on contributions (MPCR) was either 0.3 or 0.75. They found that group size *per se* did not significantly affect individuals' average contributions, although for the same MPCR contributions were if anything higher in a larger group. The levels and patterns of average contributions in their 10 person groups with MPCR = .3 are not very different from those in our 16 person baseline groups with MPCR = .2.

[10] In standard VCM experiments, subjects learn only the group's aggregate contribution, not the contributions of each individual. We made this information available in our baseline game because it must be given in the expulsion games and its availability in both conditions therefore eliminates a nonessential difference between them. The same procedure was used in baseline sessions by Bochet *et al*. and Page *et al*.

[11] A decision screen headed "you are now in period __ of 15" made sure that subjects could keep track of the number of periods remaining.

In the expulsion treatment, a second stage followed the allocation decision in each period but the last. In this stage, subjects were shown the same information about one another's current and past contributions and earnings as in the baseline condition, again in a random order and with fixed letter identifiers. Next to the information for each subject was a box that could be checked by a click of the mouse to register a vote for removing that individual from the original group, which the instructions identified as "the Green group." After each such vote, the list was shown again, with the boxes now showing the number of votes cast to expel each individual. If half or more of the current members of the Green group voted to expel a group member, that subject was moved for the remainder of the 15 periods to a secondary group dubbed "the Blue group," where the same allocation game was played, except with endowments of only five experimental dollars each period, and with no voting stage. That is, for a subject l belonging to the Blue group, period earnings were

$$(5 - C_l) + (0.2)\Sigma_{\text{all } k} C_k \tag{2}$$

where C_l is what l assigned (contributed) to the group account, and $\Sigma_{\text{all } k} C_k$ is the sum of contributions to the group account by all members of that group, l included. When the coalition to expel an individual was large enough to be effective, each subject who voted for that person's expulsion was debited 25 experimental cents (approximately 1 real cent). If the required half of current Green group members was not achieved, no money was deducted for the votes cast. Since members saw the number of votes against each individual, votes could have a warning effect even when insufficient to expel.

If the Blue group remained small, the potential earnings of its members were markedly lower than for those in the Green group. For example, if there were three subjects in the Blue group, the maximum earning of a Blue group member was 5 experimental dollars per period, while the thirteen subjects remaining in the Green group could each have earned E$26 per period.[12] Note that expulsion from the Green group was both irreversible and costly, hence the threat to expel a low contributor was a type of trigger mechanism. As with the baseline treatment, the expulsion treatment continued for

---

[12] That is, assuming all members contributed E$10. Even with universal defection, a Green group member earned E$10. If all except himself contributed E$10 while he contributed zero, a Green group member could earn E$34 ($1.19) per period.

fifteen periods, with the number of periods being announced at the outset and the current period number being displayed to subjects in each period.

We conducted eight sessions, each with sixteen inexperienced participants, for a total of 128 subjects.[13] Equal numbers of sessions were devoted to each of two different designs, each consisting of two 15-period treatments. The BE design was comprised of a (15-period) baseline treatment followed by a (15-period) expulsion treatment, allowing within-subject and within-session comparisons of behaviors with and without an expulsion option. The EE design was comprised of a first 15-period expulsion treatment followed by another 15-period expulsion treatment, to investigate learning effects.[14] There were four sessions of each type. With four sessions of each design, we have a total of sixteen session-level treatments, twelve of which were expulsion treatments, four of which were baseline treatments. Table 1 summarizes the structure of the experiment as a whole.

To prevent the subjects from anticipating the second treatment of the session and having this affect behavior in the first treatment, we told the subjects at the beginning of the session about the first one only. After the first treatment ended, we told the subjects there would be a second treatment, and that this would be the last part of the experiment. No communication was allowed in either design. Subject instructions, which were read aloud and also appeared on each subject's computer monitor, are included in the Appendix.[15]

---

[13] Subjects were drawn from the general undergraduate population at Brown University. 57% were females, 52% had never taken an economics course, and 84% were concentrators in disciplines other than economics. Each subject participated in only one session. Because subjects were drawn from a large general student population (about 5800), there were unlikely to know one another. The subjects were seated at computer terminals in a large room and were not permitted to communicate with one another during the experiment. Because the Blue group was given a task paralleling that of the Green group and because removal to the Blue group involved no physical relocation, other participants could not tell which specific individual was "expelled." Cash pay-out at the end of the experiment session was in closed envelopes. Concern about public shaming of those expelled from the Green group should accordingly have been minimal.

[14] We conducted no EB sessions because our main concern was to contrast B with E behaviors, and the BE order had training advantages. Also, we had limited resources and wanted to assure a minimal number of sessions of each treatment.

[15] The instructions shown are those used before the first treatment of the EE sessions. Prior to the second EE treatment, subjects were simply informed that they would play a second fifteen rounds by the same rules as before. The instructions read before the baseline treatment of BE sessions made no reference to expulsion or voting, so those parts of the instructions were introduced in a second instruction period before beginning the second treatment.

Rationale and Expectations

Our experiments were designed so that standard theory makes a simple prediction. We made expelling a group member costly, with both a small cost to cast an effective expulsion vote, and a (possibly much larger) loss of potential future contributions from the expelled individual. Accordingly, the expulsion option would pose no threat among payoff maximizing players fully informed of one another's preferences, and since the game had a known end-point, contributions to the public good should in theory have been at the Nash equilibrium level of zero throughout. If expulsion turned out to be an effective threat in actuality, this would suggest that most subjects believed that some might be willing to punish others at a cost to themselves. Without some preference additional to payoff maximization, or at least the belief that players having such preferences could be present, effectiveness of the expulsion threat is difficult to explain.

In the baseline condition, efficiency is monotonically increasing in contributions, as shown by equation (1). Total earnings are maximized, and full efficiency is achieved, when all subjects contribute their entire E$10 to the group account, in which case each earns E$32. But the iterated dominant strategy for payoff maximizing players is to contribute nothing, in which case each earns E$10.

In the expulsion condition, the threat of expulsion was intended to be an effective deterrent to free riding if it were believed that low contributors would be expelled and if Green group members were expected to achieve a degree of cooperation. However, the design assured that the threat of expulsion would *not* be credible if all players were payoff maximizers and knew one another to be so.[16]

Suppose, however, that subjects do not know one another's preferences or types. In particular, suppose that subjects believe that there is a non-zero probability that some players have a preference for penalizing free riders (Fehr and Gächter, 2000b, 2002), a

---

[16] That this is so in our design follows from backward induction. In the final period, there is no danger of expulsion, so all pay-off maximizing subjects can be expected to contribute zero to the group account. Since a Green group member will thus earn E$10 in the final period in any case, it would never pay for her to spend E$0.25 on expelling another member at the end of period 14. But if there is no credible expulsion threat at the end of period 14, the only equilibrium in period 14 itself is one of zero contributions. And so forth.

preference sufficiently strong to make them incur a small cost to vote for their expulsion, even after period 14.[17] The expected earnings of player i decline when some player, say j, is removed from the group, provided that j's expected total contribution to the group account during the remaining periods of play is not zero. However, this foregone earning cost as well as the voting cost of expelling j might be recouped if the expulsion of j leads *other* subjects to increase their contributions to the group account by more than the amount lost to the group by j's removal. Even a rational payoff maximizer might then calculate that it is beneficial to participate in a vote for expulsion. In fact, it could be the case that all players are payoff maximizers, but so long as enough believe others to entertain the possibility that some have a taste for punishment, it can be equilibrium behavior for actual expulsions to take place (compare the logic in Kreps *et al*., 1982). The modest direct cost of voting, only about 1 cent of real money for a period's first vote, implies that even a small subjective probability that others find tastes for punishment a credible possibility, a small amount of direct utility from punishing, or even a bit of miscalculation about optimal strategy, could be enough to trigger expulsion votes.

3. Results on Contributions and Earnings

Question 1. What was the pattern of contributions in the baseline treatment?

Result 1. *There were substantial contributions to the group account in the baseline treatment, but they declined rapidly with repetition.*

The middle line of Figure 1 shows average contributions by period in the baseline treatment of the BE sessions. Contributions in period 1 averaged 67% of endowment, but their level dropped to 31% in period 3 and continued to decline, averaging less than 20%

---

[17] Fehr and Gächter treat the preference for punishing free riders as the negative aspect of reciprocity, the punishment itself as "altruistic punishment." For other sources in the growing literature on reciprocity, see Hoffman *et al*. (1998), Ben-Ner and Putterman (2000), and Gintis *et al*., eds. (forthcoming). Our experimental design is similar to the model of Bowles and Gintis (2004), which offers a scenario by which the propensity to punish could have evolved in human pre-history. Note that positive reciprocity too (Fehr and Gächter, 2000b) could play a role in our experiment: some subjects might prefer to cooperate provided they are assured that others are doing so, so contributing itself could be a manifestation of reciprocity.

throughout the last seven periods and standing at only 8% in period 15.[18] This is much like baseline behavior in other VCM experiments.

Question 2. In expulsion treatments, how often if at all did expulsions occur?

Result 2. *In most treatments with expulsion in most sessions, a majority of subjects voted to expel another subject at least once, and in almost all expulsion treatments and sessions, one or more subjects were expelled from the Green group.*

Table 2 summarizes the average number of subjects expelled in each period of each expulsion treatment, averaged over the relevant four sessions. It shows that more than three subjects were expelled over the course of a typical expulsion treatment in BE sessions, that more than two typically suffered this fate in the first expulsion treatment of an EE session, and more than one in the second expulsion treatment of an EE session. At the session level, one or more subjects were usually expelled from the Green group in the very first vote, and one or more subjects were expelled over the course of each session of every treatment except the second treatment of one of the four EE sessions.

Question 3. What was the pattern of contributions in the Green and Blue groups in expulsion treatments? Was the threat of expulsion sufficient to raise contributions in the Green group? Overall?

Result 3a. *In expulsion treatments, contributions to the group account were higher than in baseline treatments even in the first period, indicating an anticipation effect, and they rose, rather than fell, with repetition and evidence of expulsions and expulsion votes.*

Consider first Green group contributions in the BE expulsion treatment, shown in the upper line of Figure 1. The average proportion of the endowment contributed to the public good began at 83% compared with 67% in period 1 of the baseline treatment,

---

[18]   The downward trend in contributions is confirmed by an OLS regression in which average contribution is the dependent variable and period number and a constant are the only independent variables. We include one observation per period for each of the 64 BE subjects, for a total of 960 observations. Regressions were estimated using OLS, using fixed effects for individuals, and by the Newey-West adjusted standard error procedure. In all models, the coefficient on the period term is negative and significant at the .001 level.

suggesting that subjects adjusted their behaviors in anticipation of the threat of expulsion even before a single subject had actually been expelled.[19] Contributions rose with repetition for the first twelve periods, approaching 98% of endowment in period twelve.[20]

Next, consider the pattern of contributions in the EE sessions. Figure 2 shows that the patterns in both treatments of the EE sessions were similar to that in the expulsion treatment of the BE sessions. In the first EE expulsion treatment (second line from above), Green group contributions began at 87% of endowment, rose in period 2, and remained between 90% and 94% of endowment through period fourteen.[21] In the second EE expulsion treatment (the upper line), contributions began still higher, at 94.8%, and rose to fully 100% in periods ten through thirteen before declining slightly in period fourteen.[22]

Result 3b. *Average contributions were significantly higher in expulsion treatments than in baseline treatments, even when Blue group members are included in the calculation.*

At the individual level, a within-subject comparison shows that 51 of the 52 BE subjects who were never expelled from the Green group contributed more in the treatment with expulsion than in the baseline treatment, with the remaining subject

---

[19] A within-subject comparison of first period contributions shows that of the 42 subjects whose contribution levels differed in the two treatments, 29 contributed more in the expulsion treatment. A Wilcoxin matched pair test shows the first contributions in the expulsion treatment to be higher at the 1% significance level ($z = 2.66$). Comparisons were also done between the average first period contribution in the BE baseline treatment and that in the first EE expulsion treatment and the second EE expulsion treatment, respectively, using Mann-Whitney tests with one observation for each session, hence four observations for each treatment. Both the first EE expulsion treatment and the second EE expulsion treatment were found to have higher first period contributions than the BE baseline treatment, significant at the 5% level in one-tailed tests.

[20] The significance of the upward trend is confirmed by regression estimates. We used the 723 individual-level observations for each individual and each period he or she was in the Green group, up to and including period 13. Making the individual's contribution the dependent variable and the period number the sole explanatory variable, we find a positive coefficient on period that is significant at the .001 level regardless of whether the regression is estimated using an OLS model, a model with individual fixed effects, or a model with Newey-West adjusted standard errors.

[21] Regressions again show a significant upward trend through period 13. In this case, there are 749 individual-level observations for each individual and each period he or she was in the Green group, up to that period. We find a positive coefficient on period that is significant at the 5% level using an OLS model and one significant at the 10% in a model with adjusted standard errors.

[22] Regressions here too support a significant upward trend, this time even if period 14 is included. There are 857 individual-level observations for each individual and each period he or she was in the Green group up to and including that period. We find positive coefficients on period that are significant at the 1% level using an OLS or an individual fixed effects model, while the coefficient on period is positive and significant at the 5% level using a model with adjusted standard errors.

contributing E\$10 every period. At the session level, a Mann-Whitney test comparing average contributions in the expulsion and baseline treatments of the four BE sessions, with one observation per session and treatment (hence, 8 observations in total), finds that average contributions were significantly higher with expulsion, with a p-value of .0145 in a one-tailed test, even when contributions of expelled subjects are included. Mann-Whitney tests also show contributions to be higher in each of the two EE expulsion treatments than in the BE baseline treatment, with $p = .0145$ even when the contributions of Blue group members are included.[23]

Result 3c. *Contributions declined sharply in period 15 of all three expulsion treatments.*

Although end-game effects are common in finitely repeated public goods experiments, the scale and rapidity of the period 15 declines in our expulsion treatments are dramatic. The *prima facie* case that the high levels of contributions to the group account in earlier periods were a response to an expulsion threat is bolstered by the fact that when the threat was removed in period 15 (there being no expulsion opportunity at that stage) contributions "collapsed."

Here, it is interesting to compare mechanisms that achieve cooperation through trust or reciprocity to the mechanism of the present experiment, which fosters cooperation through fear of expulsion. Groups in Bochet *et al.*'s face-to-face and chat room communication treatments without punishment, in Brosig *et al.*'s (2003) face-to-face and audio-visual conference communication treatments, and groups self-selected for cooperation in Page *et al.*'s endogenous grouping treatment, achieved levels of cooperation from first to penultimate periods that were comparable to Green group contributions in the expulsion treatments of the present experiment. In the last period, however, average contributions in those experiments declined much more modestly. It seems that while the expulsion threat in our experiment may have required the presence or at least the belief in preferences for punishing, a form of negative reciprocity, the

---

[23] In fact, average contributions, whether or not the Blue group is included, were higher in each and every one of the twelve plays of an expulsion game (that is, the BE expulsion games of four sessions, the first EE expulsion games of four sessions, and the second EE expulsion games of the same four sessions) than in any of the four plays of the baseline game.

reliance on threats crowded out the development of a more positive spirit of reciprocity among our subjects.[24]

Result 3d. *The members of the expelled group were never numerous enough to benefit from contributing to their group account.*

The lower lines in Figures 1 and 2 plot the average contribution levels of Blue group members, who had only five experimental dollars to allocate each period. Because their earnings equation, (2), contains the same factor $m = 0.2$ as does equation (1) of the Green group, it was impossible for Blue group members to benefit from contributing to their group account unless they had at least six cooperating members. But the Blue group never reached six or more members in any session. There was thus no rationale for contributing apart from altruism or confusion, and the low contributions shown on the graph are largely consistent with this. In combination with the impact of the expulsion threat on Green group contributions, this meant that the gap between Green group and Blue group earnings was usually on the order of 5-to-1, or $E20 (in real money, $0.70 per period; see Figures 3 and 4), so that a credible expulsion threat could indeed be effective in deterring free riding.

Question 4. How did learning affect behaviors in the EE sessions, when subjects engaged in a fifteen period treatment with expulsion option twice in the same session?

Result 4a. C*ontributions were higher from the outset and rose to higher levels in the second as compared to the first treatment of EE sessions, even though fewer subjects were expelled from the Green group.*

Compare the upper two lines in Figure 2, the higher of which represents Green group contributions in the second expulsion treatment of the EE sessions. The now-experienced subjects began this part of their sessions with higher contributions, evidently to avoid attracting expulsion votes.[25] Thereafter, it took higher and higher contributions

---

[24] Compare Fehr and Fischbacher's (forthcoming) remark that "explicit incentives may cause a hostile atmosphere of threat and distrust, which reduces any reciprocity-based extra effort" (see also Fehr and Rockenbach, 2003).

[25] Of the 64 subjects in the four EE sessions, 38 contributed the same amount, usually 10, in the first periods of both treatments. Of the 26 who contributed different amounts, 18 contributed more in the

to be safe from such votes, since someone contributing 8 of 10 experimental dollars was more likely to receive expulsion votes when the average was 9.5 than when it is 8.5 or 8 (we discuss voting further in section 4). Of the 52 subjects who were never expelled in the EE sessions, 10 contributed E$10 in every period of both treatments, 40 contributed more in the second than in the first treatment, and only 2 contributed less in the second than in the first treatment.[26,27] Averaged over the four sessions, as in the figure, contributions in the second treatment were higher than those in the first in every period except period 15, indicating that the expulsion threat was more effective the second time around. Averaging over all periods and including both Blue and Green group members, contributions were higher in the second than the first treatment in every session.

Result 4b. *There was a stronger end-game effect in the second than in the first treatment of EE sessions, but conclusions about backward unraveling cannot be drawn from one repetition.*

Figure 2 shows that contributions were lower in the last period of the second treatment than in that of the first treatment of the EE sessions. In every one of the four EE sessions, the size of the decline in average contribution from period 14 to period 15 was larger in the second treatment than in the first.[28] In addition, there was a small decline in contributions in period 14 of the second treatment, suggesting anticipation of the period 15 collapse. However, there is a larger period 14 decline in the expulsion treatment of the BE sessions, where there was no learning from a prior expulsion treatment. Also, despite the slight decline, period 14 contributions were higher in the second EE expulsion treatment than in either the first EE expulsion treatment or the BE

---

second treatment. A Wilcoxon matched pair test finds the first and second game first contributions to differ with a *p*-value better than 1% (z = 2.27).

[26] The Wilcoxon matched pair test statistic for the 42 subjects whose average contributions differed in the two games is significant at the 0.001 level.

[27]Table 2 shows that there were fewer subjects expelled from the Green group in the second than in the first treatment of the EE sessions. Of the four EE sessions, one had only one subject expelled in period one of the second treatment, one had no expulsions until period 14, and one had no second treatment expulsions at all, although it still supported contributions at high levels. Evidently, the threat of expulsion was made sufficiently credible in the first treatment, and was reinforced enough by individuals' votes against low contributors in the second treatment, so that many subjects needed to see no further expulsions to be persuaded to make high contributions.

[28] In three of the four sessions, the decline was more than 25% larger in the second as compared to the first game; in the remaining session, the difference was 6%.

expulsion treatment. Experimental economists find that backward induction takes a long time to learn, so we should not try to infer too much from a single repitition.

Question 5. Did Green group members *earn* more in the expulsion than in baseline treatments, and if so, was their gain enough to compensate those who lost out by being placed in the Blue group?

Result 5. *Although subjects voted out of the Green group earned far less than others, earnings averaged over all subjects (including Blue group members) were nonetheless higher in expulsion than in baseline treatments. Thus, the addition of an expulsion option increased overall efficiency, not only earnings in the non-expelled group.*

Average earnings by period and group are graphed in Figures 3 and 4. Averaging earnings across the full fifteen periods of each treatment but treating the treatments of different sessions as separate observations, we performed Mann-Whitney tests paralleling those for contributions. The differences in earnings were found to be significant to the same degree as were those in contributions, with one-tailed Mann-Whitney tests yielding *p*-values of .0145 in all three comparisons (BE expulsion versus baseline treatment, first EE expulsion treatment versus BE baseline treatment, and second EE expulsion treatment versus BE baseline treatment), even when Green group and Blue group members are given equal weight. Average earnings in every one of the twelve session-level treatments with expulsion exceeded average earnings in every one of the four BE session baseline treatments.

Question 6: How did earnings in the second EE expulsion treatment compare to those in the first?

Result 6. *Average earnings were higher with repetition, that is, higher in the second treatment of the EE sessions than in the first.*

In all four EE sessions, the average earnings, calculated over all subjects including those expelled to the Blue group, were higher in the second treatment. If average earnings in the second treatment of each of the four sesions are compared to average earnings in the first treatments in a Mann-Whitney test (which ignores which

session an observation comes from), the *p*-value for the one-sided hypothesis that earnings were higher in the second treatment is .057.

4. Voting's Incidence, Cost and Impact

The right to vote to expel a member from the Green group was used relatively sparingly. On an average of 73% of the occasions in which a subject could vote, he or she did not vote to expel anyone, while 13.5% of the time, a subject cast only one expulsion vote. An average of 22% of subjects cast no votes to expel during the course of a treatment, while 8% cast only one such vote, 24% cast between two and four expulsion votes in total per treatment, and 45% cast five or more expulsion votes in a treatment.[29]

Question 7: How costly was it to expel a member from the Green group?

Result 7. *Expulsions were costly in terms of projected lost contributions.*

The infrequency of votes to expel suggests that subjects took its costs seriously. While the perceived cost of voting a subject out of the group could have been as small as E\$0.25 (about one cent of real money) if the person targeted was not expected to contribute to the group account in any future period, observed behaviors make it unlikely that subjects usually held the latter expectation. On average, expelled subjects had contributed E\$4.82 per period during the periods prior to their expulsion, E\$2.82 during their last period before expulsion.[30] The average expelled Green group member would have had about seven more periods to contribute had she remained in the group. Had she continued to lower her contributions at the observed average rate (among those ultimately expelled) of E\$0.20 per period, she would have contributed another E\$14.14, yielding (0.2xE\$14.14=) E\$2.83 to each continuing Green group member. On this calculation, the member voting her out was depriving himself of E\$2.83, or about 10 cents real money,

---

[29] Numbers do not total 100% due to rounding error.
[30] Of thirty-two overall expulsion events in the three expulsion games, seventeen of the individuals targeted, or just over half, had contributed a positive amount in their last period in the Green group. The average last contribution of these sixteen was E\$6, or 60% of the maximum contribution possible.

for each individual expelled.[31] Of course, subjects are likely to have reasoned, we think correctly, that by expelling a few, they would induce those remaining to increase their contributions by more than enough to make up for this foregone income. Despite this, the immediate costliness of expulsions means that expulsions should not have occurred among strictly payoff maximizing agents with common knowledge of one another's preferences, so these expulsions are evidence that most subjects were either negative reciprocators, believed that others believed such preferences possibly present, or made calculating errors.[32]

Question 8: How were expulsion votes targeted?

Result 8. *Expulsion votes were targeted selectively against the lowest contributors to the Green group's group account.*

Our analysis shows that every subject expelled from the Green group in any treatment with expulsion was either the lowest or the next-to-lowest contributor in their group during their last period before expulsion. Moreover, next-to-lowest contributors were only expelled (a) when the lowest contributor was also expelled, or (b) when the next-to-lowest had a lower average contribution in previous periods as a whole.

Analysis of individuals' votes also shows them to be overwhelmingly targeted at low contributors. Several kinds of regression equations indicate that the number of votes that an individual received to be expelled was significantly negatively related to that individuals current and past contributions.

---

[31] The estimate for the rate of decline, E$0.20, was obtained by estimating a linear regression using all contributions to the Green group by those ultimately expelled in the BE expulsion treatment, with the period as independent variable. Similar results were obtained for the EE expulsion treatments, although for the first EE expulsion treatment the coefficient on period is smaller and not statistically significant. To be sure, the targeted subject might have raised, not lowered, her future contributions had she received almost but not quite enough votes to be expelled, thanks to the warning effect demonstrated in Table 4. However, in the absence of communication, there was no way for group members to coordinate so as to achieve vote totals nearing but not exceeding a half of the group's size.

[32] The willingness to engage in costly punishment when it cannot possibly have a signaling effect (here, in period 14) seems to reflect the same type of preference that is manifested when subjects engage in costly punishment of free riders in the last period of partner-group VCM's with punishment (for example, Fehr and Gächter, 2000a, Page *et al*., 2003), or in strictly one-shot perfect stranger VCM's with punishment (for example, Fehr and Gächter's perfect stranger treatment and Anderson and Putterman, 2003). See also Andreoni *et al*. (2003).

First, we estimated forty-two regression equations, one for each voting period and expulsion treatment, in which the dependent variable is the number of expulsion votes against a given subject, *j*, and the independent variables represent the information available to others about *j*'s current and past allocation choices at the time of the vote in question. The results (see Tables A1 – A3) show that subject *j* received more expulsion votes the lower were his or her contributions relative to the session average, with this difference being statistically significant either for the current period's difference or for that of the previous period, or both, in all but two of the 42 regressions estimated.[33]

Second, we estimated three OLS regressions, one for the BE expulsion treatment, one each for the first and for the second EE expulsion treatments, pooling the data for the different periods and sessions. In these regressions, the dependent variable is the number of votes to expel each individual, and the independent variables are the rank of that individual's contribution within the session and group in that period, the absolute level of his or her contribution, and a constant. In these regressions, the coefficients on both contribution rank and absolute contribution are negative and significant at the 5% level, in all but one case at the 0.1% level.

Finally, we estimated a variant of each of the last three regressions in which both the current and the previous period's contribution rank and absolute contribution are included as explanatory variables. With this specification, current contribution is negatively related to votes received, significant at the 0.1% level, for all three treatments. Contribution rank shows an equally significant negative effect in the first EE and in the BE expulsion treatment, lagged rank is also significant and negative in the first EE expulsion treatment and lagged absolute contribution is significant and negative in the second EE expulsion treatment.[34]

---

[33] In this set of specifications, we estimated a separate regression for each period because subjects had different information at different times. For example, at the time of the first expulsion vote, subjects saw only one another's first contributions; at the ninth expulsion vote, they saw separately one another's ninth, eight, and seventh contributions, and their average contributions so far. Each regression combines the data from four sessions. We used the deviation of subject *j*'s contribution from the Green group's average contribution for that period in *j*'s session, rather than the absolute level of his contribution in that period, to control for session effects. The explanatory power of these regressions tends to decline after the initial periods because there were few expulsion votes between the early periods and the last few periods of the treatment. In some regressions for later periods of the second EE expulsion treatment, where every subject contributed all ten experimental dollars period after period, the regression could not be estimated without dropping one or more variables.

[34] The regressions discussed in this and in the previous paragraph are not shown, to conserve space.

An interesting feature of this experiment is that whereas somewhere between a fifth and a third of all targeted monetary punishments in Fehr and Gächter (2000), Bochet *et al*. (forthcoming) and Page *et al*. (2004) were aimed perversely at groups' *highest* contributors,[35] perverse punishment never occurred in the present experiment. The contrast between the presence of perverse monetary punishment and the absence of perverse punishment by expulsion can be attributed to the fact that the desire to punish high contributors, often as revenge for previous punishment by them, is almost never exhibited by more than a minority of subjects, and it can thus be nullified by putting control over discipline in the hands of the majority. Indeed, about 18% of votes to expel in the present experiment were cast against contributors of more than the current session average, but not a single high contributor was ever expelled.[36]

Question 9: Is there evidence of a successful signaling effect on remaining Green group members after an expulsion?

Result 9. *Successful expulsions triggered upward adjustments in Green group contribution levels.*

We estimated a tobit regression (see Table 3) at the session level in which the change in the average contribution of members remaining in the Green group from period t to period $t + 1$ is the dependent variable the number of members voted out of the Green group during the vote following period t's contribution decisions is the main explanatory variable of interest.[37] Other independent variables, included as controls, are the period number (because contributions tended to increase less as the treatment proceeded), dummy variables for treatment (BE expulsion treatment, etc.), and in one of the two

---

[35] Based on our calculations from the raw data, including data provided by Fehr and Gächter.

[36] Ertan et al. (in progress) report on a public goods experiment in which subjects could impose costly punishment on other group members after learning of their contributions, but only if a majority in their group voted to allow punishment over the range of contributions in question. In 160 votes, no group ever voted to allow punishment of contributors of more than the average amount to the public good. Both findings illustrate that with those inclined to punish perversely being in the minority, it is efficiency-enhancing to have the power to limit punishment be in the hands of the majority.

[37] There is one observation per pair of periods for the expulsion game of each BE session, one for each pair of periods for the first expulsion game of each EE session, and one for each pair of periods for the second expulsion game of each EE session. The change from period 14 to period 15 is left out because in period 15, the expulsion threat was entirely absent. The regression was estimated in Stata with the cnreg command, treating observations in which the average contribution of period $t+1$ is 0 or 10 as being potentially censored.

specifications shown, the average contribution of the Green group members in period t. The result is a highly significant coefficient on the number expelled, suggesting that the average contribution rose by E\$0.86 after a successful expulsion. The change in contributions is significantly negatively related to the period number.

Question 10: Is there evidence of a successful signaling effect on subjects who received some but not sufficient votes to be expelled?

Result 10: *The more votes to expel a subject, the more that subject's contribution increased in the next period if the subject was not expelled from the Green group. The impact of votes declined as the end of the treatment approached.*

For each expulsion treatment, we estimated two regression equations, each with one observation for each Green group member and pair of periods.[38] In the first regression model shown for each treatment, in Table 4, the percentage change in subject j's contribution from period t to period $t + 1$ is the dependent variable, and the independent variables are (a) a constrant, (b) the number of votes to expel subject j, (c) j's contribution in period t, and (d) the period number. The second regression model of each pair is the same as the first except that the product (or interaction) of the period number and the number of votes to expul j is included as an additional explanatory variable, to see whether the impact of expulsion votes weakened as the end of the treatment approached. The coefficient on the number of expulsion votes received is positive in the estimates for all three parts, and is significant at the .001 level in all estimates except in the model without interaction term for the second EE expulsion treatment. The estimates without interaction terms indicate that on average, one additional expulsion vote caused subject j to increase her contribution by 3.2% in the BE expulsion treatment, and by 4.4% in the first EE expulsion treatment. The coefficients on contribution are significant and negative, presumably because the room for upward adjustment was smaller the more j was already contributing. In the second model, the coefficient on the interaction term is negative and significant for the two EE treatments, indicating that the impact of votes on

---

[38] For each subject remaining in the Green group for the entire treatment, the regression includes thirteen observations, the change from period 14 to 15 again being left out (see the previous note). Observations for subjects who were eventually expelled from the Green group are also included up to their expulsion; thus, for a subject playing in the Green group for periods 1 to 10 only, there would be nine observations.

the change in contribution was declining with the period, as would be expected with rational subjects facing a threat of a punishment the cost of which falls with time.[39]

Question 11: Is there a correlation between individuals' contributions and their votes to expel? In particular, were high contributors more likely to cast expulsion votes than were low contributors?

Result 11. *The correlation between contributing and voting to expel low contributors is significant and negative.*

For each subject, we averaged the number of votes he or she cast against contributors of less than the Green group period average and the amount contributed to the group account during all periods the individual was in the Green group. In all three treatments, the bivariate correlation between these two variables is negative and significant at the 0.01% level. If the same correlation is calculated including only those individuals who were never expelled from the Green group, it is still negative, but it is no longer statistically significant except in the BE expulsion treatment, where it is significant at the 5% level.

One might have conjectured that, since both contributing to the group account and voting to expel free-riders are civic acts that foster group well-being, the high contributors would have voted more frequently to expel low contributors.   Our results instead suggest that it was the more reluctant contributors who were more likely to do so, perhaps with the mind-set that "if I'm going to be dragged into making these contributions, against my own interest, I'm not going to put up with others' free riding." It seems that among our subjects, the tendency to punish is not highly correlated with the tendency to contribute to the group account.[40]

5. Conclusion

---

[39] A referee pointed out that some subjects who received expulsion votes may have revised their contributions upwards not simply because this seemed prudent, given the danger of being expelled, but also because while less than a majority of votes did not translate into a monetary punishment, it was read as a stern expression of disapproval, as in Masclet *et al*. (2003).

[40] Page et al. (2004) report a similar result: in their treatment combining endogenous group formation and punishment opportunities, they find the most punishment in their low contributor groups.

Like a trigger strategy, the expulsion mechanism threatens a dire consequence with irreversibility.[41] In the experiment, probably much of the effectiveness of the expulsion mechanism is due to the fact that expulsion was used so sparingly, less than three times on average in a treatment of 15 periods. Low contributors were often warned by increasing but still less-than-majority votes before they were expelled, and most of the low contributors responded to the increased warnings with higher contributions, thus avoiding expulsion. The "dark side" of the mechanism is that by relying so heavily on the implied threat of expulsion, it didn't allow group members to develop trust in one another's positive reciprocity, so cooperation collapsed when the threat was removed.

In expulsion, punishment, and regrouping mechanisms, a Bayesian interpretation suggests two tendencies that work in opposite directions. Often individuals attempt to signal and build a reputation for being a high contributor (or cooperator) early in the game, and then cash in on their reputations by decreasing their contributions toward the end of the game. With increased experience and learning, both tendencies may strengthen at the same time. In Figure 2, the second expulsion game has higher contributions in the early periods than in the first expulsion treatment; but at the same time the end-game fall-off is larger. Guttman (2003) predicts that these two tendencies will be in balance even with fully rational and fully experienced players for a repeated prisoner's dilemma game, and McKelvey and Palfrey (1991) predict the two tendencies will be in balance both in theory and in experiment, for the centipede game.

What implications do our results have for real world economic and social institutions? In partnerships, communities, and organizations of various types, people often strive to cooperate in the absence of state enforcement. The threat that an uncooperative individual may be ostracized or expelled is one way to foster cooperation. If, in the experiment or the larger world, each individual cares only for her own material outcome, without emotions of annoyance at others' free riding and without values of cooperation or reciprocity, this threat would not be credible. Our results suggest that human nature includes other types of preferences as well as self-interested preferences and that this heterogeneity renders credible what otherwise is not a credible threat, making the danger of exclusion a reason to pull one's weight in a group endeavor.

---

[41] Some trigger strategies may be reversed eventually; see Fudenberg and Tirole, 1991, p. 185.

References

Ahn, T.K., Elinor Ostrom and James Walker, Feb. 2002, "Incorporating Motivational Heterogeneity into Game Theoretic Models of Collective Action," unpublished paper, Workshop on Political Theory, Indiana University.

Alchian, Armen and Harold Demsetz, 1972, "Production, Information Costs, and Economic Organization," *American Economic Review* 62: 777-95.

Anderson, Christopher and Louis Putterman, 2003, "Do Non-strategic Sanctions Obey the Law of Demand? The Demand for Punishment in the Voluntary Contribution Mechanism," Brown University Department of Economics Working Paper 2003-15.

Andreoni, James, 1988, "Why Free Ride? Strategies and Learning in Public Goods Experiments," *J. of Public Economics* 37: 291-304.

Andreoni, James, William Harbaugh, and Lise Vesterlund, 2003, "The Carrot or the Stick: Rewards, Punishment and Cooperation," *American Economic Review* 93 (3): 893-902.

Ben-Ner, Avner and Louis Putterman, 2000, "On Some Implications of Evolutionary Psychology for the Study of Preferences and Institutions," *Journal of Economic Behavior and Organization* 43: 91-99.

Bochet, Olivier, Talbot Page and Louis Putterman, 2002, "Communication and Punishment in Voluntary Contribution Experiments," Working Paper No. 2002-29, Brown University Department of Economics. (Available at http://www.econ.brown.edu/ 2002/)

Bowles, Samuel and Herbert Gintis, 2004, "The Evolution of Strong Reciprocity: Cooperation in Heterogeneous Groups," *Theoretical Population Biology* 65: 17-28.

Brosig, Jeannette, Joachim Weimann and Axel Ockenfels, 2003, "The Effect of Communication Media on Cooperation?" *German Economic Review* 4 (2): 217-41.

Carpenter, Jeffrey and Peter Matthews, 2002, "Social reciprocity," Middlebury College Department of Economics Working Paper #29.

Casari, Marco and Charles R. Plott, 2003, "Decentralized Management of Common Property Resrouces: Experiments with a Centuries-Old Institution," *Journal of Economic Behavior and Organization* 51 (2): 217-47.

Davis, Douglas D. and Charles A. Holt, 1993, *Experimental Economics.* Princeton: Princeton University Press.

Djankow, Simeon and Peter Murrell, 2002, "Enterprise Restructuring in Transition: A Quantitative Survey," *Journal of Economic Literature* 40 (3): 739-92.

Ehrhart, Karl-Martin and Claudia Keser, 1999, "Mobility and Cooperation: On the Run," CIRANO 99s-24, Montreal,

Ertan, Arhan, Talbot Page and Louis Putterman, in process, "Public Choice and Private Sanctions in a Public Goods Game: An Experiment with Endogenous Institutions," unpublished paper, Brown University Department of Economics.

Fehr, Ernst and Simon Gächter, 2000a, "Cooperation and Punishment," *American Economic Review* 90: 980-94.

Fehr, Ernst and Simon Gächter, 2000b, "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives* 14: 159-81.

Fehr, Ernst and Simon Gächter, 2002, "Altruistic Punishment in Humans," *Nature* 415: 137-40.

Fehr, Ernst and Urs Fischbacher, forthcoming, "The Economics of Reciprocity," in H. Gintis, S. Bowles, R. Boyd and E. Fehr, eds., *Moral Sentiments and Material Interests: The Foundations of Cooperation In Economic Life*. Cambridge: MIT Press.

Fehr, Ernst and Bettina Rockenbach, 2003, "Detrimental Effects of Sanctions on Human Altruism," *Nature* 422: 137-140.

Fischbacher, Urs, Simon Gächter and Ernst Fehr, 2001, "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment," *Economics Letters* 71: 397-404.

Frank, Robert, 1988, *Passions within Reason*. New York: W. W. Norton.

Frohlich, Norman and Joe Oppenheimer, 1998, "Some Consequences of e-mail vs. Face-to-Face Communication in Experiment," *Journal of Economic Behavior and Organization* 35 (3): 389-403.

Fudenberg, Drew and Jean Tirole, 1991, *Game Theory*. Cambridge, MA: MIT Press.

Gintis, Herbert, Samuel Bowles, Robert Boyd and Ernst Fehr, eds., forthcoming, *Moral Sentiments and Material Interests: The Foundations of Cooperation In Economic Life*. Cambridge: MIT Press.

Gunnthorsdottir, Anna, Daniel Houser, Kevin McCabe, and Holly Ameden, 2002, "Disposition, History and Contributions in a Public Goods Experiment," unpublished manuscript, Department of Economics and Economic Science Laboratory, University of Arizona.

Guttman, Joel, 2003, "Repeated Interaction and the Evolution of Preferences for Reciprocity," *Economic Journal* 113 (489): 631-56.

Hirshleifer, David and Eric Rasmusen, 1989, "Cooperation in a Repeated Prisoners Dilemma with Ostracism," *Journal of Economic Behavior and Organization* 12 (1): 87-106.

Hirshleifer, Jack, 1987, "On the Emotions as Guarantors of Threats and Promises," in John Dupré, *The Latest on the Best: Essays on Evolution and Optimality*. Cambrdige: MIT Press.

Hoffman, Elizabeth, Kevin McCabe and Vernon Smith, 1998, "Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology," *Economic Inquiry* 36: 335-52.

Isaac, R. Mark and James M. Walker, 1988a, "Communication and Free-Riding Behavior: The Voluntary Contributions Mechanism," *Economic Inquiry* 26: 585-608.

Isaac, R. Mark and James Walker, 1988b, "Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism," *Quarterly Journal of Economics* 103 (1): 179-99.

Kreps, David, Paul Milgrom, John Roberts and Robert Wilson, 1982, "Rational Cooperation in Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory* 27: 245-52.

Kurzban, Robert and Daniel Houser, 2001, "Individual Differences in Cooperation in a Circular Public Goods Game," *European Journal of Personality* 15 (S1): S37-S52.

Masclet, David, Charles Noussair, Steven Tucker and Marie-Claire Villeval, 2003, "Monetary and Non-Monetary Punishment in the Voluntary Contributions Mechanism," *American Economic Review* 93 (1): 366-80.

McKelvey, Richard and Thomas Palfrey, 1992, "An Experimental Study of the Centipede Game," *Econometrica* 60 (4): 803-36.

Offerman, Theo, Joep Sonnemans and Arthur Schram, 1996, "Value Orientations, Expectations, and Voluntary Contributions in Public Goods," *Economic Journal* 106: 817-45.

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

Page, Talbot, Louis Putterman and Bulent Unel, 2004, "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry, and Efficiency," Working Paper 2002-19, Department of Economics, Brown University, revised May, 2004.

Sally, David, 1995, "Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992," *Rationality and Society* 7 (1): 58-92.

Schelling, Thomas, 1960, *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Sefton, Martin, Robert Shupp and James Walker, 2002, "The Effect of Rewards and Sanctions in Provision of Public Goods," Working Paper, University of Nottingham and Indiana University.

Shapiro, Carl and Joseph Stiglitz, 1984, "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review* 74: 433-44.

Table 1. Summary of Designs and Treatments

Design	Treatment #	Description	Number of Sessions	Total No. of Subjects
BE	1	Baseline treatment, 15 periods.	4	64
	2	Expulsion treatment, 15 periods.		
EE	1	Expulsion treatment, 15 periods.	4	64
	2	Expulsion treatment, 15 periods.		
Experiment as A Whole			8	128

Table 2. Cumulative number of subjects expelled from Green group at the end of the period (equivalent to average size of Blue group in following period).

<div align="center">

Treatment

</div>

Period	BE expulsion treatment	EE, 1st treatment	EE, 2nd treatment
1	2	1	0.75
2	2	1	0.75
3	2	1	0.75
4	2.25	1	0.75
5	2.25	1.5	0.75
6	2.25	2	0.75
7	2.25	2	0.75
8	2.25	2	0.75
9	2.5	2.25	0.75
10	2.5	2.25	0.75
11	2.5	2.25	0.75
12	2.5	2.25	0.75
13	3	2.25	0.75
14	4.5	2.25	1.25

Note: each entry is an average over four sessions.

Table 3: Tobit regression estimates of change in average contribution in Green group from period t to period $t + 1$ as a function of the number of members expelled at the end of period t, the period number, and the average contribution in the session in period t.

Variables	Without average contribution ($\overline{C_t}$). Coefficients	With average Contribution ($\overline{C_t}$). Coefficients
$\overline{C_t}$		0.1216**
		(0.0541)
expelled	0.3091***	0 .3732***
	(0.0891)	(0.0942)
BE1 dummy	-0.2524**	-0.2170*
	(0.1123)	(0.1150)
EE1 dummy	-0.3225***	-0.2627**
	(0.1104)	(0.1140)
t	-0.0259**	-0.0289**
	(0.0116)	(0.0118)
Constant	0.5454 ***	-0.6110
	(0.1224)	0.5240
# of Obs.	156	156
Prob>chi^2	0.0000	0.0000
PseudoR^2	13	15

Notes: The regression includes one observation per session, expulsion treatment and pair of adjacent periods, 1,2,...14. The dependent variable is the change in the average contribution to the Green group account between period t and period $t + 1$, taken over all individuals belonging to a session and treatment's Green group in period $t + 1$. Independent variables: ***Expelled*** = number expelled (voted out) from Green group during the vote that follows the contributions of period t; $\overline{C_t}$: = average contribution in the Green group in period t by members remaining in the Green group in period t+1; ***BE dummy*** = 1 for observations from the BE expulsion treatment, 0 otherwise; ***EE(1) dummy*** = 1 for observations from the first EE expulsion treatment, 0 otherwise (omitted category is second EE expulsion treatment); t = period number of the base period of the observed change.

Table 4: **Regression estimates of percentage change in contribution as a function of the number of expulsion votes received, contribution, period, and a period*votes interaction term.**

Variables	Treatment					
	BE expulsion treatment		EE, 1st treatment		EE, 2nd treatment	
	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
V_{ji}	2.723***	2.682 **	4.129***	5.848***	0.501	2.526 ***
	(0.613)	(1.198)	(0.454)	(0.944)	(0.322)	(0.819)
C_{jt}	-9.231***	-9.247***	-10.334***	-9.890***	-4.893***	-3.863***
	(0.954)	(1.032)	(0.542)	(0.581)	(0.752)	(0.841)
$V_{ji}*t$		0.005		-0.207**		-0.200***
		(0.125)		(0.099)		(0.074)
t	-0.249**	-0.251*	-0.196*	-0.082	-0.047	0.003
	(0.127)	(0.134)	(0.105)	(0.118)	(0.071)	(0.073)
Constant	88.247***	88.410***	95.816***	90.747***	48.148***	37.427***
	(8.952)	(9.849)	(5.249)	(5.778)	(7.282)	(8.277)
# of Obs.	676	676	715	715	793	793
R^2	0.163	0.163	0.340	0.352	0.089	0.100

Notes: All regressions include individual fixed effects, not shown. Dependent variable: $(C_{i,t+1} - C_{it})/ C_{it}$. Independent variables: V_{ji} = number of votes cast by any subject j to expel subject i; C_{it} = contribution by subject i in period t; $V_{ji}*t$ = interaction term between the number of votes and the period number. Numbers in parentheses are standard errors. *** = significant at the 1% level; ** = significant at the 5% level and *=significant at the 10% level.
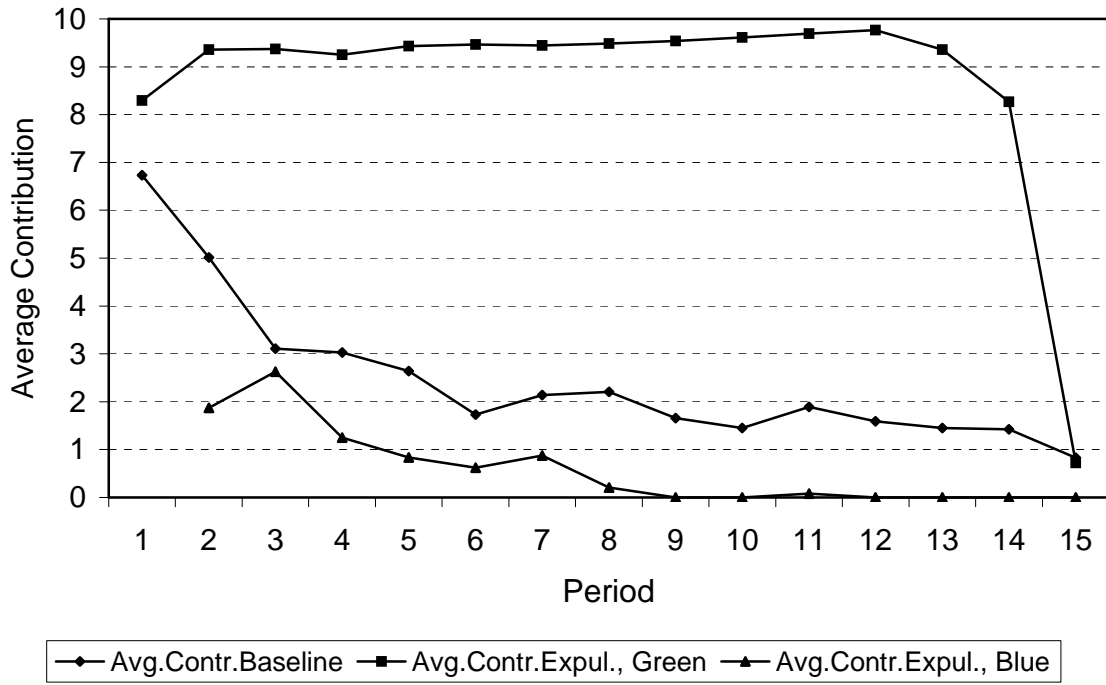
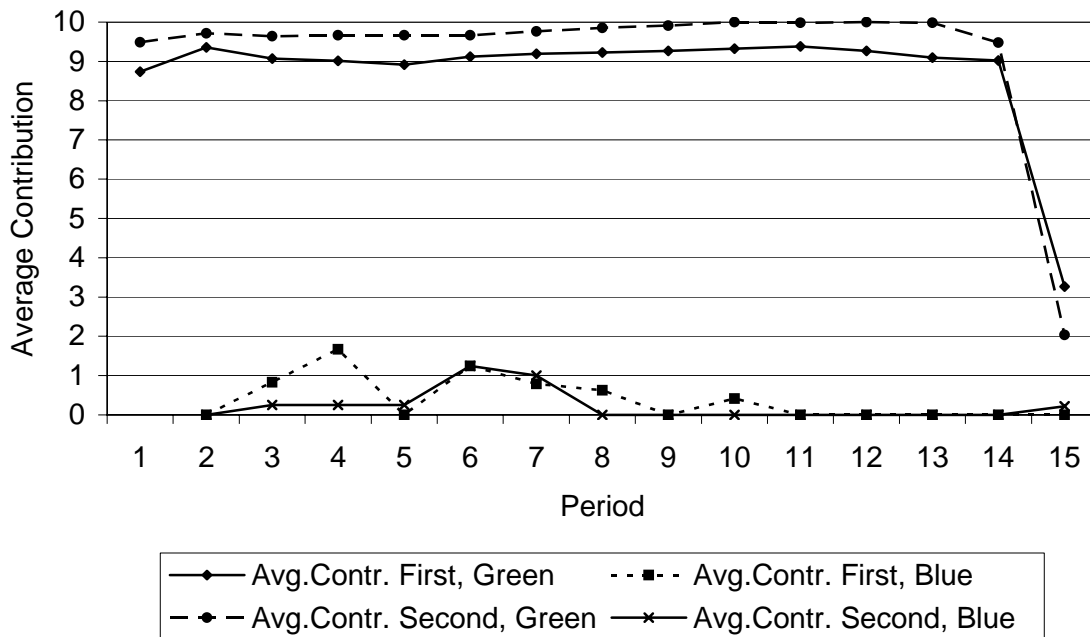Figure 1. Average contribution in BE design, by treatment and group.

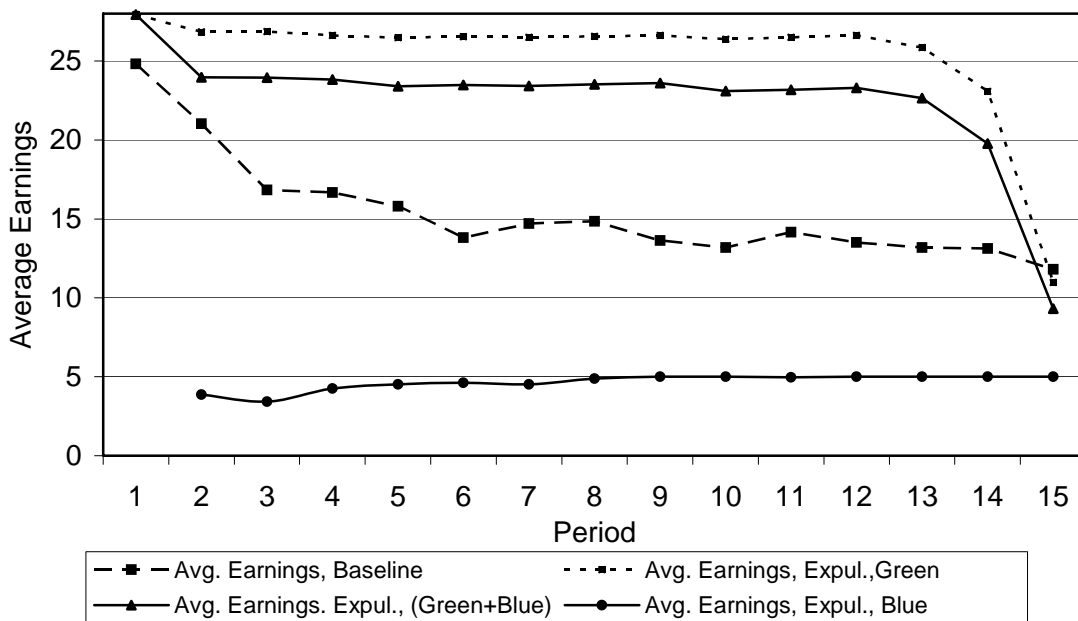Figure 2. Average Contribution in EE Design by Treatment and Group.

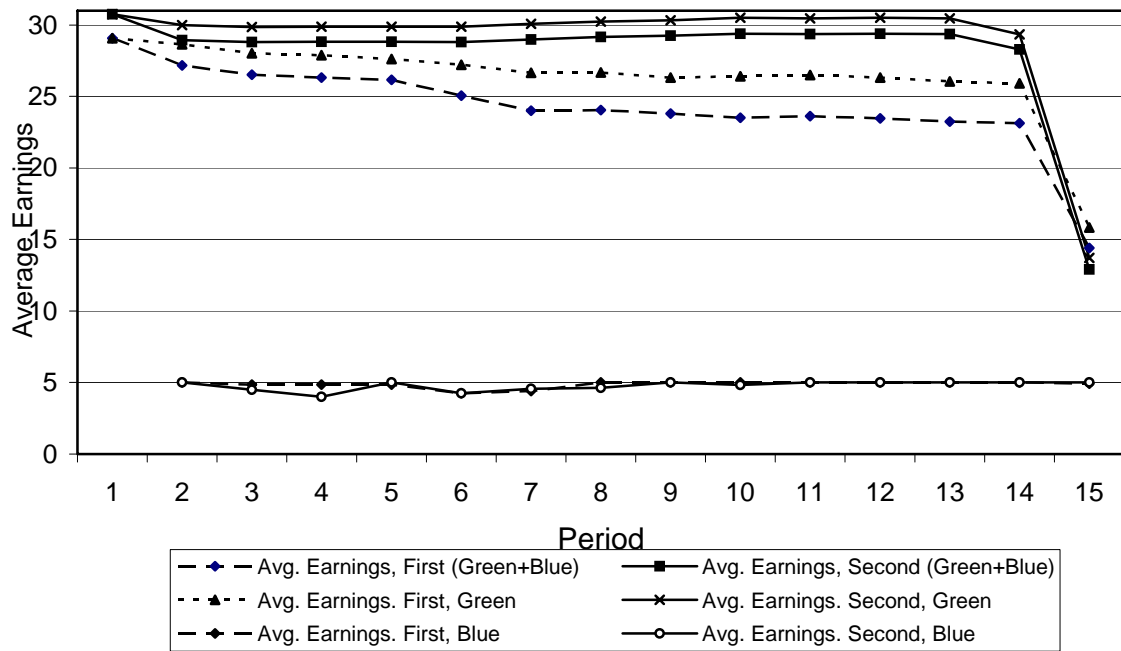Figure 3. Average Earnings in BE Design by Treatment and Group.

Figure 4. Average earnings in EE design by treatment and group.

Appendix

Instructions for Expulsion Treatment

[Screen 1]
This is an experiment, funded by a research foundation, to study decision making. You will be earning money in "experimental dollars" during the experiment. At the end of the experiment you will be paid in cash in real dollars (each experimental dollar is worth a real $0.035, that is three and a half cents). The amount you will earn will depend on your and others' decisions. The maximum possible earning is $42 (real dollars) and the minimum possible is $5. You are likely to earn an amount in between. Please make sure you understand the decision process.

[Screen 2]
Rounds and Groups

There are 15 rounds (or periods) in the experiment. There are two groups in the experiment: a **Green Group** and a **Blue Group**. At the beginning of the experiment all the participants start in the Green Group and no one is in the Blue Group. What happens in the Green Group differs somewhat from what happens in the Blue Group, and we will explain what happens in the Green Group first.

[Screen 3]

Instructions for Members of the Green Group – the **Allocation Decision**

At the beginning of each round, each member of the Green group receives $10 (experimental dollars). As a member of the group, you must decide how to divide this amount between the Green group account and your personal account. Each of the others in your group also receives $10 (experimental dollars), which he or she allocates between the group account and his or her personal account. After you make your decision and everyone else makes their decision, everyone will learn the total amount that has been allocated to the group account.

The money you assign to your personal account goes into your earnings.

An amount equal to 0.2 times the sum of the allocations to the Green group account goes into your earnings.

Earnings = (amount in personal account) + (0.2)(total in Green group account)

[Screen 4]

Examples of Earnings Calculation

To understand how your earnings are affected by the different allocation decisions that might be made, we consider a few examples.

Example 1. There are 16 people in the Green group and each one, including you, put all ten dollars into the group account.

The total in the group account would be $10x16 = $160.

The amount in your personal account would be $10 – amount put in group account = $0.

Earnings = (amount in personal account) + (0.2)(total in Green group account)

= 0 + (0.2)($160) = 0 + $32 = $32

The next screen shows how you enter your decision. Go to this screen, type the number 10, press enter, and you will see how your earnings would be calculated under this first scenario.

[Practice Screen 1 here]

[Screen 5]

Example II. There are 16 people in the Green group and each one, including you, put zero dollars into the group account.

The total in the group account would be $0x16 = $0.

The amount in your personal account would be $10 – amount put in the group account = $10.

Earnings = (amount in personal account) + (0.2)(total in Green group account)

= $10 + (0.2)($0) = $10 + $0 = $10

Go to the next screen, type the number 0, press enter, and you will see how your earnings would be calculated under this second scenario.

[Screen 6]

Practice Questions

During the experiment, each individual makes his or her allocation decision before learning what the others decisions are. Different group members may make different decisions. To help you further your understanding of how these decisions affect your earnings, we will now work through two additional practice screens. For each screen, you

will find a paper worksheet on your desk. We will fill out the first worksheet together, then the first screen.

Practice Worksheet 1

Fill out the worksheet for this example. You are in the Green group and there are 16 people in the group. Each has $10 to allocate.

       Each of 8 of the others allocates $0 to the group account
       Each of 7 of the others allocates $10 to the group account
       You allocate $10 to the group account.

1. Amount you allocate to the group account: _____10

2. Amount you allocate to your personal account: _____0
       ($10 – your allocation to the group account on line 1)

3. Total number of dollars allocated to the group account: _____80

4. Income from the group account for a member of your group   _____16
       (0.2)(group account total in line 3)

5. Your earnings from this round: _____16
       (group account income in line 4 + personal account in line 2)

Now, go back to the practice screen. Type in your allocation to the group account and press enter to make sure your calculations are correct.

[Practice Screen 3]

Practice Worksheet 2

This is the similar to Worksheet 1.

       Each of 8 of the others allocates $0 to the group account
       Each of 7 of the others allocates $10 to the group account
       You allocate $0 to the group account.

1. Amount you allocate to group account: _____0

2. Amount you allocate to your personal account: _____10
       ($10 – your allocation to the group account on line 1)

3. Total number of dollars allocated to the group account: _____70

4. Income from the group account for a member of your group   _____14
       (0.2)(group account total in line 3)

5. Your earnings from this round: _____24
       (group account income in line 4 + personal account in line 2)

Now, go back to the practice screen. Type in your allocation to the group account and press enter to make sure your calculations are correct.

[Practice Screen 4]

[Screen 7]

Instructions for Members of the Green Group – the **Voting Decision**

In the second part of each round, you will get information on the others' allocations to the group account. You will see on your screen a list that contains for each person in the Green group:

    (1) his or her average amount allocated to the group account so far in the experiment,
    (2) the amounts he or she allocated to the group account in a few of the most recent rounds, and
    (3) his or her earnings in the current round and accumulated earnings in the experiment so far (in experimental dollars)

The list will be in a random order, except that information for members of the Blue group, if any, will appear at the bottom, in blue. Also, for your reference, your average allocation and a few of the most recent allocations will be included in the list. Your entry will have "you" printed next to it. The other members of the group will be identified to you by letters B, C, D, etc. These letters will be randomly assigned at the start of the experiment and each will refer to the same person for all 15 rounds. You will not learn which actual person corresponds to each letter either during the experiment or afterwards.

Just to the right of the information for each member of your group, you will find a box. Clicking inside the box produces a check mark which will constitute a vote to remove the individual in question and place him or her in the Blue group. You can change your mind by clicking again inside the box before you submit your votes. You can vote to remove as many or a few as you want. When you finish voting, click in the box labeled "Finished Voting" at the bottom of the screen.

[Screen 8]

Results of the Green Group's Voting

It costs something for members of the Green group to remove a fellow member to the Blue group. Specifically, when a vote to remove someone from the group attains the necessary number of votes, each person voting to remove the individual in question will be charged 25 experimental cents, which will be deducted from their overall earnings in the experiment. If there are not enough votes to remove an individual, there is no charge to those who voted to do so.

If half or more of the current members of the Green group vote to remove a particular person, the person is moved from the Green group to the Blue group. Once a person is moved to the Blue group, this person remains in the Blue group for the rest of the experiment. Remember that during the experiment and after it is over, you will not know the actual identity of this person, or any actual identity of any person in either group.

You will see the results of the voting on the screen. The number in each box shows you how many votes were cast to move that particular person to the Blue group. In your own box you will see the number of votes cast against you. Those who will be moved to the Blue group will be shown in blue.

[Screen 9]
Instructions for Members of the Blue Group

If you have been moved to the Blue group you will stay in this group for the rest of the experiment. You have no voting decision. In each round you will have an allocation decision. This decision is the same as the allocation decision for the Green group, but you will have $5 rather than $10 to allocate.

As before, the money you assign to your personal account goes into your earnings.

Also as before, an amount equal to 0.2 times the sum of the allocations to the Blue group account goes into your earnings.

Earnings = (amount in personal account) + (0.2)(total in Blue group account)

To make sure that you understand how earnings in the Blue group are determined, you have been given two practice worksheets, labeled Practice Worksheet 3 and Practice Worksheet 4. Please fill out each sheet in the same manner as before and then type the indicated allocation in the corresponding practice screen to check your results. Raise your hand if you have questions.

Practice Worksheet 3

Fill out the worksheet for this example. You are in the Blue group and there are 4 people in it. Each has $5 to allocate. Every member allocates $5 to the group account and $0 to his or her personal account.

1. Amount you allocate to the group account: _____5

2. Amount you allocate to your personal account: _____0
    (= $10 - group account allocation on line 1)

3. **Total** number of dollars allocated to the group account: _____20

4. Income from the group account for each member of the group _____4
    (0.2)(group account total in line 3)

5. Your earnings from this round: _____4
    (group account income in line 4 + personal account income in line 2)

Now, go back to the practice screen. Type in your assignment to the group account according to the scenario above and press enter to make sure your calculations are correct.

[Practice Screen 5]

Practice Worksheet 4

Fill out the worksheet for this example. You are in the Blue group and there are 6 people in it. Each has $5 to allocate. Every member allocates $5 to the group account and $0 to his or her personal account.

1. Amount you allocate to the group account: _____5

2. Amount you allocate to your personal account: _____0
    (= $10 - group account allocation on line 1)

3. **Total** number of dollars allocated to the group account: _____30

4. Income from the group account for each member of the group _____6
    (0.2)(group account total in line 3)

5. Your earnings from this round: _____6
    (group account income in line 4 + personal account income in line 2)

Now, go back to the practice screen. Type in your assignment to the group according to the scenario above and press enter to make sure your calculations are correct.

[Practice Screen 6]

[Screen 9]

While members of the Green group are being shown one another's allocations and making their voting decisions, members of the Blue group will simply see a listing of one another's allocations. Although they do not vote, they will have to wait until the first group completes their voting before proceeding with their next allocation stage.

[Screen 10]

Earnings and Group Size

Note that the only people who can allocate to the Blue group account are members of the Blue group, and likewise for the Green group. Earnings in one group are independent of those in the other. The size of the group matters, and the Blue group is likely to start out small. Size also matters in the Green group. As we have seen, for example, when there are 16 members and each puts $10 into the group account, each earns $32. If there are only 12 members and each puts $10 into the group account, each earns $120(0.2) = $24.

[Screen 11]

Your Payment in Real Dollars

At the end of round 15, your earnings for the experiment will be totaled and converted from experimental to real dollars, which will be paid to you in cash. If your earnings are less than $5 (real dollars), you will be paid $5 (real dollars).

[Screen 12]

During the experiment, you are asked to be as quiet as possible. There is to be no communication other than the information transmitted via the computers. Any violation of this rule may result in your dismissal from the experiment and forfeiting your earnings. Also please note that typographical errors cannot be corrected once decisions have been submitted, so be careful that you have entered your choices accurately before pressing the enter key.

Any questions?

Table A.1: **OLS regressions: number of votes to expel _j_ as a function of the difference between _j_'s contribution and session Green group averages in current and past periods, BE expulsion treatment.**

Period(t)	# Obs. and R^2	$C_{j,t} - \overline{C_t}$	$C_{j,t} - \overline{C_{t-1}}$	$C_{j,t} - \overline{C_{t-2}}$	$(1/(t-3))\sum_{s=1}^{t-3}\left\{C_{j,s} - \overline{C_s}\right\}$
1	64 **0.71**	-0.887** (0.073)	n.a.	n.a	n.a.
2	60 **0.57**	-0.729** (0.085)	0.019 (0.049)	n.a.	n.a.
3	60 **0.59**	-0.496** (0.124)	-0.522** (0.107)	-0.106 (0.058)	n.a.
4	60 **0.53**	-0.841** (0.165)	-0.054 (0.172)	-0.471** (0.158)	-0.022 (0.080)
5	60 **0.50**	-1.307** (0.217)	-0.068 (0.182)	-0.240 (0.178)	-0.177 (0.166)
6	58 **0.40**	-1.296** (0.228)	-0.379 (0.226)	0.180 (0.188)	0.289 (0.188)
7	56 **0.24**	-0.850* (0.247)	0.198 (0.239)	-0.171 (0.221)	0.079 (0.231)
8	56 **0.12**	-0.555 (0.389)	-0.312 (0.260)	-0.073 (0.259)	-0.130 (0.258)
9	55 **0.65**	-1.794** (0.207)	-0.121 (0.189)	0.247 (0.136)	0.143 (0.152)
10	55 **0.28**	-1.237** (0.299)	-0.142 (0.215)	-0.240 (0.198)	-0.041 (0.192)
11	55 **0.29**	-0.820** (0.232)	0.435 (0.354)	-0.309 (0.268)	-0.241 (0.221)
12	55 **0.13**	-0.720* (0.302)	0.091 (0.245)	-0.285 (0.366)	-0.204 (0.280)
13	55 **0.15**	-1.161* (0.505)	-0.250 (0.344)	-0.188 (0.283)	0.065 (0.359)
14	55 **0.11**	-1.008* (0.407)	-0.195 (0.518)	-0.247 (0.321)	-0.004 (0.402)

Notes: Numbers in parentheses are standard errors. The right-hand column is the average from periods 1 to (_t_-3) of the difference between _j_'s contribution and the Green Group average in that period. n.a. means no coefficient is estimated because the variable does not exist in this period.   ** = significant at the 1% level; *= significant at the 5% level. Column 2 from the left contains the sample size and the R-square.

$\overline{C_t} = \sum_{all'i'} C_{i,t}$, where i indexes Green group members; j included.

Table A.2: OLS regressions: number of votes to expel j as a function of the difference between j's contribution and the session Green group average in current and past periods, first EE treatment.

Period(t)	# Obs. and R^2	$C_{j,t} - \overline{C_t}$	$C_{j,t} - \overline{C_{t-1}}$	$C_{j,t} - \overline{C_{t-2}}$	$(1/(t-3))\sum_{s=1}^{t-3}\left\{C_{j,s} - \overline{C_s}\right\}$
1	64 **0.87**	-1.206** (0.058)	n.a.	n.a	n.a.
2	56 **0.45**	-0.737** (0.112)	0.010 (0.038)	n.a.	n.a.
3	56 **0.51**	-0.994** (0.195)	-0.131 (0.148)	-0.005 (0.039)	n.a.
4	56 **0.28**	1.002** (0.283)	0.013 (0.204)	0.015 (0.172)	-0.049 (0.041)
5	55 **0.33**	-1.254** (0.256)	0.102 (0.256)	0.097 (0.179)	-0.053 (0.078)
6	55 **0.40**	-1.528** (0.327)	0.498 (0.360)	-0.387 (0.247)	-0.061 (0.106)
7	55 **0.37**	-1.509** (0.329)	-0.105 (0.313)	0.230 (0.351)	0.05076 (0.114)
8	55 **0.38**	-1.082** (0.275)	0.582 (0.299)	-0.192 (0.247)	-0.019 (0.139)
9	55 **0.41**	-2.396** (0.455)	0.509 (0.399)	-1.608** (0.480)	0.076 (0.242)
10	54 **0.23**	-0.933** (0.265)	0.010 (0.213)	-0.035 (0.210)	-0.166 (0.146)
11	54 **0.49**	-2.530** (0.384)	0.459 (0.406)	1.021** (0.356)	-0.132 (0.256)
12	54 **0.43**	-1.487** (0.290)	-0.302 (0.228)	0.532 (0.282)	0.071 (0.193)
13	54 **0.80**	-0.855** (0.067)	0.708 (0.433)	-0.952** (0.332)	-0.083 (0.324)
14	52 **0.09**	-0.075* (0.037)	0.010 (0.062)	0.160 (0.386)	-0.069 (0.335)

See notes to Table A.1.

Table A.3: **OLS regressions: number of votes to expel *j* as a function of the difference between *j*'s contribution and the session Green group averages in current and periods, second EE treatment.**

Period(t)	# Obs. and R^2	$C_{j,t} - \overline{C_t}$	$C_{j,t} - \overline{C_{t-1}}$	$C_{j,t} - \overline{C_{t-2}}$	$(1/(t-3))\sum_{s=1}^{t-3}\left\{C_{j,s} - \overline{C_s}\right\}$
1	64 0.45	-1.568** (0.220)	n.a.	n.a	n.a.
2	61 0.56	-1.287** (0.152)	-0.077 (0.081)	n.a.	n.a.
3	61 0.35	-1.429** (0.329)	-0.286 (0.239)	-0.110 (0.118)	n.a.
4	61 0.29	-1.726** (0.389)	0.449 (0.433)	0.092 (0.318)	-0.152 (0.149)
5	61 0.48	-1.618** (0.332)	-0.468 (0.333)	0.039 (0.384)	-0.303 (0.226)
6	61 0.41	-2.776** (0.540)	0.019 (0.420)	-0.076 (0.419)	-0.171 (0.416)
7	61 0.17	1.065 (1.038)	-0.852* (0.356)	-0.045 (0.268)	-0.621 (0.377)
8	61 0.48	-2.515** (0.606)	1.968 (1.325)	-0.448 (0.459)	-1.166 (0.551)
9	61 0.45	-4.222** (0.836)	0.616 (0.945)	1.581 (1.616)	-1.279 (0.854)
10	61 0.03	dropped	-0.413 (0.453)	0.698 (0.507)	-0.311 (0.395)
11	61 0.51	-5.397** (0.673)	dropped	-0.060 (0.397)	-0.285 (0.386)
12	61 0.35	dropped	-3.463** (0.623)	dropped	-0.281 (0.349)
13	61 0.19	-3.693** (1.210)	dropped	-2.693* (1.210)	0.002 (0.753)
14	61 0.67	-0.684** (0.063)	0.232 (1.065)	dropped	-0.448 (0.724)

See notes to Table A.1. Some independent variables were dropped by the estimating program due to high correlations resulting from lack of change in contributions across these periods.