

# An Experiment on Nash Implementation \*

Antonio Cabrales<sup>†</sup>  
Gary Charness<sup>‡</sup>  
Luis C. Corchón<sup>§</sup>

June 22, 1998

---

\*We would like to thank James Costain, Rosemarie Nagel, Diego Rodríguez and a seminar audience at Universitat Pompeu Fabra for helpful comments. The financial support of Spain's Ministry of Education under grant PB96-0302 and Generalitat de Catalunya under grant 1997SGR-00138 is gratefully acknowledged.

<sup>†</sup>Universitat Pompeu Fabra. Ramon Trias Fargas 25. E-08005 Barcelona. Spain.

<sup>‡</sup>Universitat Pompeu Fabra

<sup>§</sup>Universidad Carlos III. Madrid 126. E-28903 Getafe. Spain.

## Abstract

We perform an experimental test of Maskin's canonical mechanism for Nash implementation, using 3 subjects in non-repeated groups, as well as 3 outcomes, states of nature, and integer choices. We find that this mechanism successfully implements the desired outcome a large majority of the time and an imbedded comprehension test indicates that subjects were generally able to comprehend their decision tasks. The performance can also be improved by imposing a fine on non-designated dissidents. We offer some explanations for the imperfect implementation, including risk preferences, the possibilities that agents have for collusion, and the mixed strategy equilibria of the game.

*Key Words:* Implementation; Experiments; Mechanisms.

*JEL classification:* C72, C92, D70, D78.

# 1 Introduction

A standard economic problem involves a central planner trying to implement an optimal outcome. Often the planner has only highly constrained access to information about true preferences. The theory of implementation addresses the problem of designing mechanisms whose equilibria<sup>1</sup> satisfy certain socially desirable properties, but which do not require that the authorities have unrealistically accurate information about the underlying parameters of the economy. The role of the authorities is simply to ascertain that the rules of the game are respected by all participants and to implement the outcome decreed by the mechanism.

Mechanism design is particularly important in economic environments with public goods, as it is well known that a voluntary contributions mechanism will not implement a social choice function which tries to achieve Pareto efficient outcomes. Suboptimal provision is commonly found in experiments (Isaac and Walker 1988, e.g.), although the degree of deficiency of the voluntary contributions is not as high as predicted by neo-classical theory. Since this mechanism is not able to achieve desired outcomes, it is important to study alternatives.

Chen and Tang (1996) have done a comparative study of the basic quadratic mechanism of Groves and Ledyard (1977) and the Paired-Difference Mechanism by Walker (1981). The Groves and Ledyard (1977) mechanism is only effective in a very constrained set of environments (quasilinear preferences, linear production technologies), while the Walker (1981) mechanism can handle more general environments, but can only implement the Lindahl correspondence<sup>2</sup>. We study experimentally the canonical mechanism for implementation in Nash equilibria (see Maskin 1977, Repullo 1987). This mechanism can implement a wider variety of social choice rules, under a much larger domain of preferences, than the other mechanisms mentioned. Other mechanisms, like the one proposed by Abreu and Matsushima (1992), can implement an even wider variety of social choice functions (although the implementation is “virtual”, i.e. with arbitrarily high probability). But this mechanism has been criticised by Glaser and Rosenthal (1992) and this criticism has been supported experimentally by Sefton and Yavaş (1995). In addition, Cabrales (1997) has shown that if the game defined by the mechanism and the agents’ preferences is played repeatedly by boundedly rational agents, the process converges to the desired (and stable) solution.<sup>3</sup> Yet the mechanism is quite controversial. According to Jackson (1992) “A nagging criticism of the theory is that the mechanisms used in

---

<sup>1</sup>Or more appropriately, the equilibria of the game defined by the mechanisms and the state of the world

<sup>2</sup>Elbittar and Kagel (1997) have compared the performance of Moore’s (1992) and Perry and Reny’s (1994) mechanisms to implement the efficient allocation of an indivisible private good among two players (King Solomon’s dilemma). Both mechanisms seem to perform inadequately.

<sup>3</sup>For this to be true the traditional assumption of *monotonicity* (Maskin 1977) is replaced by *strict monotonicity*, the possible preference profiles and outcomes of the social choice rule are finite (although outcomes that are not part of the social choice rule can be infinite), and some punishments are possible.

the general constructive proofs have ‘unnatural’ features.” Moore (1992) also complains that the mechanisms for Nash implementation are “highly complex - often employing some unconvincing device such as an integer game.” Under these circumstances it seems appropriate to test the theory.

Our experiment models an environment where there are three states of the world (preference profiles) and three outcomes<sup>4</sup>. The preferences of the three agents cycle around these three outcomes in the three states, but the social choice rule picks a specific one in each state of Nature. We concentrate on such a rule for several reasons. First, we show that it is not implementable in dominant strategies, so that it is natural to look for a mechanism that implements in Nash equilibria. Furthermore, there is no focal choice in each state, and some agent does sufficiently badly (in comparative terms) under the social choice rule that telling the truth is not an obvious choice absent an enforcing mechanism. At the same time the environment is simple enough to be handled by the subjects.<sup>5</sup> This means that one of issues raised against this mechanisms -namely complexity- is avoided in our experiment. Therefore our work focuses on unnatural features -namely integer games- and the existence of mixed strategy equilibria. We will see that we find some support for the latter.

The canonical mechanism has an infinite strategy space; the agents are required to announce a state of the world, an outcome and an integer  $i \in R^+$ . If everyone announces the same state, or if there is only one *dissident* from a consensus announcement, the integer is not relevant in determining the outcome. If there is more than one dissident, the person who announces the highest number gets her announced outcome. It is easy to prove that such a game cannot have an equilibrium where, with positive probability, there is more than one dissident. The artful part of the design is to permit and require the right kind of consensus to be the equilibrium. One immediate problem is that it does not seem possible to design an experiment with a truly infinite strategy space. It is certainly not possible to allow the announcement of arbitrarily large integers and at the same time have a finite duration for the experiment. In our design we use a common modification of the mechanism in which players can choose from a finite number (3 in our case) of integers, and the winner is the player whose number is the remainder of the sum of announced integers. With this modification, the mechanism implements the social choice rule in *pure-strategy* Nash equilibrium, that is, the only pure-strategy Nash equilibrium outcome for each state corresponds to the outcome of the social choice rule. However, there are mixed strategy equilibria which have different outcomes with positive probability.

We have run 4 sessions with two treatments. The first (the baseline treatment) uses the mechanism as described. In the 2nd treatment, we introduce a *fine* if and only if there is exactly one dissident and this person is not the *designated dissident* for the state

---

<sup>4</sup>Although we focus on three outcomes, there could be infinitely many in the economy; only three are necessary for the mechanism.

<sup>5</sup>The strategy space provides a couple of ways to test the comprehension achieved by the subjects. There are good indications that the structure was well-understood by the participants.

announced by the majority.<sup>6</sup> The likelihood of the desired outcome being implemented was .68 in the baseline treatment and .80 in the 2nd treatment. The first thing to note is that there is a very substantial degree of implementation under both treatments. Additionally the introduction of fines makes the mechanism work significantly better, which indicates that there is nevertheless room for improvement to the canonical mechanism. We will argue that the improvement of the mechanism might arise because of the destabilizing effect of the fines on a mixed strategy equilibrium of the game.

## 2 The mechanism

Let us first describe the environment in which the mechanism has to be used. There are three individuals indexed by  $i \in \{1, 2, 3\}$ ; three possible outcomes:  $a, b, c$ ; and three states of the world: *red*, *yellow*, and *green*. The preferences of the individuals among the outcomes in the three states of the world can be described by:

Preferences:

Player \ state	<i>red</i>	<i>yellow</i>	<i>green</i>
1	$a \succ b \succ c$	$b \succ c \succ a$	$c \succ a \succ b$
2	$b \succ c \succ a$	$c \succ a \succ b$	$a \succ b \succ c$
3	$c \succ a \succ b$	$a \succ b \succ c$	$b \succ c \succ a$

With these preferences any deterministic single-valued social choice function must, in every state, assign the worst outcome in the preference ordering to one of the players. This will be seen to have important implications for the properties of the mechanisms that implement such social choice function.

We now introduce the social choice function that we wish to implement with our experimental design:

Social choice function:

$$F(\textit{red}) = a, \quad F(\textit{yellow}) = c, \quad F(\textit{green}) = b$$

**Proposition 1.** The social choice function  $F(\cdot)$  cannot be implemented in dominant strategies.

**Proof:** See the appendix.

---

<sup>6</sup>The dissident is a designated dissident if and only if exactly two members of a group announce a color and this color would give the dissident his highest payoff if this color were the true state and all members were to announce this color.

This result makes apparent the necessity of implementing with a more demanding equilibrium concept. The obvious choice in this case is to implement in Nash equilibrium. We will use a version of the canonical mechanism for Nash implementation (Maskin 1977, Repullo 1987, McKelvey 1989). Let us now describe the mechanism.

Strategy space: Let  $\Theta = \{r, y, g\}$  be the set of states, where  $r$  represents *red*,  $y$  represents *yellow* and  $g$  represents *green*. Let  $\Lambda = \{a, b, c\}$ , and  $N = \{1, 2, 3\}$ . The individual strategies belong to  $\Theta \times \Lambda \times N$ .

Outcome function:

1. If the 3 individuals announce:  $r$ , the outcome is  $F(\text{red}) = a$ .  
 $y$ , the outcome is  $F(\text{yellow}) = c$ .  
 $g$ , the outcome is  $F(\text{green}) = b$ .
2. If exactly two agents announce  $r$  and:  
the dissident is 1 who announces  $y$ , the outcome is  $b$ ,  
the dissident is 1 who announces  $g$ , the outcome is  $c$ ,  
otherwise the outcome is  $a$ .
3. If exactly two agents announce  $y$  and:  
the dissident is 2 who announces  $r$ , the outcome is  $b$ ,  
the dissident is 2 who announces  $g$ , the outcome is  $a$ ,  
otherwise the outcome is  $c$ .
4. If exactly two agents announce  $g$  and:  
the dissident is 3 who announces  $r$ , the outcome is  $c$ ,  
the dissident is 3 who announces  $y$ , the outcome is  $a$ ,  
otherwise the outcome is  $b$ .
5. If the three agents announce different states, then the integers announced by the three players are added up.  
If the integers add to 4 or 7 then the outcome is the one announced by player 1.  
If the integers add to 5 or 8 then the outcome is the one announced by player 2.  
If the integers add to 3, 6 or 9 then the outcome is the one announced by player 1.

**Proposition 2.** The previously described mechanism implements  $F(\cdot)$  in pure strategy Nash equilibrium.

**Proof:** See the appendix.

Notice that this mechanism implements the  $F(\cdot)$  in *pure strategy Nash equilibrium*. There are mixed strategy equilibria which produce outcomes different from the ones in  $F(\cdot)$ .

Those equilibria will be useful to understand the experimental results. The classical version of the canonical mechanism <sup>7</sup> implements in *pure and mixed* strategy Nash equilibria. The difference in the mechanism is that in the classical version the players can announce *any* integer, and the outcome is the one announced by the person who announces the highest integer (ties can be broken arbitrarily). This mechanism presents a conceptual problem, as the reason why no mixed strategy equilibrium exists is that a player may have no best response for some mixed strategy profiles of the opponents<sup>8</sup>. But, more significant from our point of view, it seems impossible to implement such a game in the laboratory. There was no evident way to actually allow players to use *any* arbitrary integer and still maintain a finite duration for the game. We could have *said* that players were allowed to use arbitrary integers, but given that time is finite, this creates constraints on the *actual* size of the integer and these constraints (like speed in writing zeros, or knowledge of shorter ways to describe numbers) are not even common knowledge. For these reasons we selected the present version of the mechanism.

Another important thing to notice is that the pure strategy equilibria of the mechanism are such that some players are using a *weak best response*. The reason is that the outcome these agents receive in equilibrium is the least preferred one for them. Thus, there would be no harm in changing the strategy used, if the other players continue using the equilibrium strategies. This, however, does not imply that the equilibrium strategy (which involves announcing the true state) is *weakly dominated* for the player who gets the least preferred outcome under  $F(\cdot)$ . It can be checked (from the table that summarizes payoffs in the instructions) that there are some combinations of strategies for the other players such that announcing the true state results in the most preferred outcome. Even taking this into account, it will be clear from the experimental results that the incentive to deviate from the equilibrium is quite important.

However, there are some weakly dominated strategies. If an agent does not announce her most preferred outcome under the true state of the world, she is using a strategy that is weakly dominated (by another that announces the same state and integer and the most preferred outcome). This will serve us as an indirect check (“rationality test”) of the degree of comprehension achieved by the agents about the working of the mechanism.

To check the importance of the fact that equilibrium strategies are weak best responses for some agents, we created a modified version of the mechanism which modifies rules 2, 3, and 4. By punishing a deviation from a dissident who is not the *designated* one. In this way the mechanism implements in *strict* Nash equilibria, since all Nash equilibria are strict. Since strict equilibria have to be in pure strategies, there is no need to emphasize the pure strategy aspect here. The complete rules for the new mechanism can be found in the instructions appendix, but we describe here the new rule 2.

## 2. If exactly two agents announce $r$ and:

---

<sup>7</sup>As exposed, for example in Repullo 1987.

<sup>8</sup>See Jackson (1992).

the dissident is 1 who announces  $y$ , the outcome is  $b$ ,  
the dissident is 1 who announces  $g$ , the outcome is  $c$ ,  
otherwise the outcome is  $a$ , and the dissident pays a fine of  $x$  pesetas.

As we will see in the data section, this change makes an important difference in the behavior of the players and the proportion of times that the outcome  $F(\cdot)$  is attained. The following corollary is straightforward from Proposition 2.

**Corollary 2.** The modified mechanism implements  $F(\cdot)$  in strict Nash equilibrium.

### 3 Experimental design

Four sessions, each with 15 participants, were conducted at Universitat Pompeu Fabra in Barcelona. The average net pay was about \$10 per subject and sessions lasted less than 2 hours.

Achieving comprehension and salience in this experiment was a challenge. At the beginning of a session, the instructions and a decision sheet were passed out to each subject. The decision sheet stated the subject number and type. Instructions covered all rules used to determine the outcome for each group and the resulting payoffs to each player in the group; these were read aloud to the entire room. As the experimental set-up is not a familiar environment, the instructions also contained an example where the states of Nature were types of weather, the outcomes were activities, and the 3 types had different state-dependent preferences among these activities.<sup>9</sup> The complete instructions can be found in the Appendix. To aid comprehension, we included complete payoff tables and seven exercise questions. These exercises were discussed aloud and questions were fielded. When the instructional phase was concluded, we proceeded with the session. As there were 5 subjects of each type, we had 5 groups of three in each of the 10 rounds of the experiment. These groups were varied - an anonymous matching process was devised so that no two groups ever had the same composition. This non-repeat feature was common knowledge.

At the start of a round, a monitor made a blind draw (with replacement) of a colored card from a box held by the experimenter. This box contained 3 yellow, 4 green, and 5 red cards.<sup>10</sup> The color drawn was the state of Nature and was known to all. On their decision sheets, participants then announced a color, an integer from  $\{1, 2, 3\}$ , and a preferred outcome. If all three members of a group announced different colors, the sum of the integers chosen determined which member's preferred outcome was implemented.<sup>11</sup> As

---

<sup>9</sup>We thank James Costain for this idea.

<sup>10</sup>This deliberate asymmetry was an attempt to create a bit of friction, perhaps making successful implementation somewhat more difficult.

<sup>11</sup>The process: if the sum of the 3 announced integers was 4 or 7, the declared outcome of the type 1



each type was aware of the true state of Nature, each type has a unique preference among the possible outcomes. Thus, we have one rationality test imbedded in the experiment - if a subject did not choose his preferred outcome, it would appear that the instructions were not well-understood.

The decision sheets were collected, announcements collated, and outcomes and payoffs determined. An individual's payoff for the period was written on his or her decision sheet (see Appendix YY) and the sheets were returned to the subjects. The next round was then initiated by another draw from the box of colored cards. Everyone was aware that the experiment would continue until 10 rounds were completed. At the end of the session, subjects were paid based on the payoffs achieved in a randomly-selected round.<sup>12</sup>

As mentioned earlier, two types of sessions were conducted. The baseline session featured payoffs of 500, 1000, or 1500 pesetas, with 500 pesetas added as a show-up fee. In the 2nd treatment, where we explore whether the disincentive of a fine would enhance the mechanism's success rate, a fine of 100 or 200 pesetas (depending on the combination of type and the state of Nature; see the instructions) was deducted from a non-designated dissident's payoff. This mechanism cannot guarantee that it is the dissident who is making a false announcement, but false reporting certainly becomes riskier.

At the end of the session, each participant was paid individually and privately.

## 4 Results

Detailed data for all sessions are shown in Appendix ZZ. We find that the social choice function was successfully implemented in 68 of 100 instances (35 of 50 in session 1; 33 of 50 in session 2) in the baseline treatment. This rate increased to 80% (39 of 50 in session 3; 41 of 50 in session 4) in the treatment with a fine for a dissident. Figure AA shows the rate of successful implementation by periods for each treatment. There is no clear trend across time. The rationality test provided by the announced preferred outcomes also indicates a reasonable level of general comprehension, as 34 of the 60 subjects selected the appropriate outcome in all 10 rounds and the mean rationality "score" was 84%. Table WW shows the distribution of rationality scores.

We also find that the proportion of subjects who announce the true state follows the consistent pattern where the likelihood of a true announcement is directly related to the payoff a subject would receive if all group members reported the state truthfully. This pattern is reassuring and provides some evidence that the subjects understood the payoffs in the game. In the baseline treatment, the likelihood of a true announcement is .93, .54,

---

person was chosen. If this sum was 5 or 8, the type 2's choice was implemented. It can be seen that this resulted in each type having the same 1/3 probability of having his or her declared outcome selected.

<sup>12</sup>This was done to make payoffs more salient to the subjects, as this method makes the nominal payoffs 10 times as large as would be the case if payoffs were aggregated over 10 periods.

and .23 for the 3 types, so that the overall rate of truth-telling is 56.7%. In the 2nd treatment, these rates by type were .90, .66, and .45, with the overall rate of honest reporting at 67.0%. Notice that the biggest behavioral change, where the associated rate of truthful announcements nearly doubles, is observed for the type who would receive the lowest payoff if all types tell the truth.

The success of the mechanism in implementing  $F(\cdot)$  is considerable. Comparing the observed baseline success rate with the expected success rate of 1/3 for a random mechanism, we find that the test of proportions gives a t-statistic of 5.22, significant at  $p \ll .01$ . The success increases substantially when the implementation is performed with our fine mechanism; the test of proportions between the two treatments gives a t-statistic of 1.95, significant at  $p < .03$  (one-tailed test). In addition to the fine increasing the success rate, we also find that the fine also reduces the sample variance for the number of successful implementations by 72%, from 2.40 in the baseline treatment to 0.67. There appears to be less uncertainty in this environment. This inference is also supported by the observation that the integer game was needed to determine the outcome 18% of the time in the baseline case, but with a likelihood of only 7% when a fine was possible.

## 5 Discussion

Two things stand out in the data. The first is that although the mechanism achieves substantial success, an outcome different from what  $F(\cdot)$  indicates is still implemented a non-negligible proportion of the time and the players who would not get their favorite outcome with universal honesty make false announcements fairly frequently. The second is that the observed behavior presents more truth telling when the fines are introduced. We will now propose some potential explanations for these findings in turn.

Let us first concentrate on the game without fines. To make the analysis simpler, we will focus the exposition on the case where the true state is *red*. An isomorphic analysis would follow for *yellow* or *green*. When the true state is *red* the type 1 agent gets her favorite outcome under  $F(\cdot)$ , type 3 gets her middle outcome and type 2 gets her worst outcome. Notice that out of 45 possible true *red* announcements in the two sessions without a fine (3 periods in the first session and 6 periods in the second, with five type 1 players in each period) the type 1 announces *red* in 43 of them. It turns out that given the actual strategies of the other types, this is indeed a best response. For this reason we need not worry much about the behavior of type 1. The behavior of type 2 and type 3 is a little harder to explain. Consider the game that results for player 2 and 3 once the strategy of player 1 is fixed. To simplify even further, assume that the three integers are used about one third of the time and that players always announce their most favorite outcome in the true state of the world (which are also in line with observed behavior). The game between types 2 and 3 is then as follows:

2\3	<i>red</i>	<i>yellow</i>	<i>green</i>
<i>red</i>	5, 10	5, 10	5, 10
<i>yellow</i>	5, 10	10, 15	10, 10
<i>green</i>	5, 10	10, 10	15, 5

In this reduced game the pair (*red*, *red*) is still an equilibrium, but notice that the strategy *red* is weakly dominated for player 2. Also, strategy *green* is weakly dominated for player 3. This explains easily the small frequency of these strategies in the data.

After the elimination of *red* for 2 and *green* for 3, we could eliminate *red* for 3. In this way we get to a component of equilibria which puts weight on *yellow* and *green* for 2 and all the weight on *yellow* for 3. Notice, however that once agent 1 is considered in the game, the only equilibrium profile in that component is the one where 1 announces *red*, 2 announces *yellow*, and 3 announces *yellow* (if 3 announces *yellow* with positive probability, 1 would like to deviate to *yellow* or *green*). If instead we eliminate *yellow* for 3, we have an equilibrium component where 2 uses *green* only, and 3 mixes between *red* and *yellow*. If the probability of *red* by 2 is high enough, agent 1 does not want to deviate. So, in principle, all of the observed behavior can more or less be accounted for with some equilibrium in one of the components we just described.

Some readers may not find this explanation entirely satisfactory. Notice that for player 3 to use *red*, she must be certain that player 2 never uses *yellow*, which is rather counterfactual if one looks at the data. However, remember that we did not tell the subjects what other players had done in the past but only the payoffs obtained, so it is possible for a subject of type 3 to maintain the belief that player 2 uses strategy *green* all the time given their information<sup>13</sup>. Similarly, for player 2 to use *yellow* she must be certain that 3 does not use *green* (this is less strong since it is not so counterfactual).

If one feels that these extreme beliefs are not very plausible, there is another way to explain the observed behavior, by taking into account that the experimental design has not controlled for risk preferences<sup>14</sup>. Consider what happens if we have a population composed of agents with heterogeneous preferences for risk. The von-Neumann Morgenstern utilities of the players for the 3 states can be described by  $(5, 10, x)$ <sup>15</sup>. The game now looks like this:

---

<sup>13</sup>Notice that player 3 can obtain a payoff of 15 even with a (*green*, *yellow*) strategy pair by winning the integer game.

<sup>14</sup>We could have used the binary lottery procedure (Roth and Malouf 1979), but there is considerable controversy about whether this actually solves the problem or makes matter worse (Selten, Sadrieh and Abbink 1995). From a practical point of view we thought that the instructions were already complicated enough without adding one more level of complexity through the explanation of the binary lottery.

<sup>15</sup>Remember that von Neumann Morgenstern utilities are invariant to affine transformations, which gives two degrees of freedom in specifying them (slope and intercept), so this description is without loss of generality.

$2 \setminus 3$	<i>red</i>	<i>yellow</i>	<i>green</i>
<i>red</i>	5, 10	5, 10	5, 10
<i>yellow</i>	5, 10	10, $x_3$	$\frac{5+10+x_2}{3}, \frac{5+10+x_3}{3}$
<i>green</i>	5, 10	$\frac{5+10+x_2}{3}, \frac{5+10+x_3}{3}$	$x_2, 5$

Where  $x_2$  is private information of player 2 and  $x_3$  is private information of player 3. Strategy *red* for 2 and *green* for 3 are still dominated, and we eliminate them. For player 2 all risk averse agents choose *yellow* and all risk seeking agents choose *green*, independently of the value of  $x_2$  as long as both *red* and *yellow* are used by a positive fraction of player 3.

The behavior of player 3 is a little harder to explain. Suppose player 3 knows that player 2 chooses *yellow* 50% of the time and *green* the remaining 50% (for this it suffices that the distribution of  $x_2$  has 50% below 15 and 50% above 15, the exact shape is immaterial). They have to compare the payoff of *red* which is 10, to the payoff of *yellow* which is  $\frac{1}{2}x_3 + \frac{1}{2}\left(5 + \frac{x_3}{3}\right)$ . All risk seeking agents prefer *yellow*. The risk averse agents prefer *red* as long as  $x_3 < \frac{45}{4}$ .

Under these conditions, we have that a game in which the value of  $x$  is lower than  $\frac{45}{4}$  for at least 50% of the population, and another 50% is risk-prone has an equilibrium where 50% of player 2 uses *yellow* and the rest uses *green*, and 50% of player 3 uses *red* and the rest *yellow*. This is qualitative similar to the observed data.

The games when the true state are *yellow* and *green* can be similarly analyzed. In the case of *yellow* player 2 gets her favorite choice and almost always announces *yellow*. The game for players 3 and 1 (who take the roles respectively of 2 and 3) is now:

$3 \setminus 1$	<i>yellow</i>	<i>green</i>	<i>red</i>
<i>yellow</i>	5, 10	5, 10	5, 10
<i>green</i>	5, 10	10, 15	10, 10
<i>red</i>	5, 10	10, 10	15, 5

It can be seen that, with the appropriate relabeling, the analysis also follows in this case.

Finally, for true state *green* we have:

$1 \setminus 2$	<i>green</i>	<i>red</i>	<i>yellow</i>
<i>green</i>	5, 10	5, 10	5, 10
<i>red</i>	5, 10	10, 15	10, 10
<i>yellow</i>	5, 10	10, 10	15, 5

The other relevant feature of the data is that the introduction of fines increases considerably the amount of truth-telling by players 2 and 3, and consequently the proportion of

times that implementation is achieved. If we again simplify the game by assuming that the true state is *red*, player 1 tells the truth and all integers are used one third of the time, the game between types 2 and 3 is now as follows:

2\3	<i>red</i>	<i>yellow</i>	<i>green</i>
<i>red</i>	5, 10	5, 8	5, 8
<i>yellow</i>	4, 10	10, 15	10, 10
<i>green</i>	4, 10	10, 10	15, 5

Now strategy *green* is still weakly dominated for player 3 (and it is still observed very infrequently). Strategy *red* is not weakly dominated any longer for player 2. The pair (*red*, *red*) is now a strict Nash equilibrium in the reduced game. Unsurprisingly, the proportion of truth telling for type 3 (the least favored one under  $F(\cdot)$ ) goes from  $6/45$  to  $13/30$ , and for type 2 (the player who gets the middle outcome under  $F(\cdot)$ ) it jumps from  $23/45$  to  $22/30$ .

The component of equilibria that we mentioned in the analysis of results in the game without fines still exist in the game with fines. We can use those equilibria (or the heterogeneity in risk preferences), as we did before, to explain the remaining deviations from implementation that are observed in the data.

## 6 Summary

We find that the canonical mechanism for Nash implementation can be quite successful in implementing the social choice function, with an observed success rate of 68% in the baseline treatment. With the inclusion of a fine for “dissidence,” the mechanism’s performance increases to 80%. While this is not perfect implementation, agents’ behavior can be better understood by taking into account possible risk preferences and the mixed strategy equilibria of the underlying game. Criticisms that such a mechanism would prove too complex seem to be unfounded here, as imbedded comprehension tests offer evidence that most participants understood the structure of the environment. A note of caution is due here, though. Our environment has only 3 states. It is conceivable that the perceived complexity will increase dramatically when the number of states increases.

On the other hand, we found evidence that mixed strategy Nash equilibria matters confirming the criticism made by Jackson (1992) to the canonical mechanism. In our case the problems caused by the mixed strategy equilibria can be significantly reduced by a simple modification of the game.

Our results suggest some further improvements for an implementation mechanism and directions for future study. It is hoped that further empirical research will lead to enhanced effectiveness in implementing desirable social choice functions.

## References

- D. Abreu and H. Matsushima (1992), "Virtual Implementation in Iteratively Undominated Strategies: Complete Information", *Econometrica*, 60, 993-1008.
- A. Cabrales (1997), "Adaptive Dynamics and the Implementation Problem with Complete Information", Universitat Pompeu Fabra Working Paper 179.
- Y. Chen and F. F. Tang, (1996) "Learning and Incentive Compatible Mechanisms for Public Goods Provision: An Experimental Study", SFB 303, Universität Bonn, Discussion Paper No. A-529.
- A. Elbittar and J. Kagel, (1997), "King Solomon's Dilemma: An Experimental Study on Implementation", mimeo, University of Pittsburgh.
- J. Glazer and R.W. Rosenthal (1992), "A Note on Abreu-Matsushima Mechanisms", *Econometrica*, 60, 1435-1438.
- T. Groves and J. Ledyard (1977), "Optimal Allocation of Public Goods: a Solution to the Free Rider Problem", *Econometrica*, 45, 783-809.
- R.M. Isaac and J. Walker (1988), "Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism", *Quarterly Journal of Economics*, 103, 179-199.
- M. O. Jackson (1992), "Implementation in Undominated Strategies: A Look at Bounded Mechanisms", *Review of Economic Studies*, 59, 757-775.
- E. Maskin (1977), "Nash Implementation and Welfare Optimality", mimeo, Massachusetts Institute of Technology.
- R.D. McKelvey (1989), "Game Forms for Nash Implementation of General Social Choice Correspondences", 6, 139-156.
- J. Moore (1992), "Implementation in Environments with Complete Information", in J. J. Laffont ed., *Advances in Economic Theory: Sixth World Congress*, Econometric Society.
- M. Perry and P.J. Reny (1994), "A General Solution to King Solomon's Dilemma", mimeo, University of Pittsburgh, forthcoming *Games and Economic Behavior*.
- R. Repullo (1987), "A Simple Proof of Maskin's Theorem on Nash Implementation", *Social Choice and Welfare*, 4, 39-41.
- A. Roth and M.W.K. Malouf (1979), "Game-Theoretic Models and the Role of Bargaining", *Psychological Review*, 86, 574-594.
- M. Sefton and A. Yavaş (1995), "Abreu-Matsushima Mechanisms: Experimental Evidence", mimeo, Manchester University.
- R. Selten, A. Sadrieh, and K. Abbink (1995), "Money does not Induce Risk Neutral Behavior, but Binary Lotteries do Even Worse", Discussion Paper B-343, Universität

Bonn, SFB 303.

M. Walker (1981), "A Simple Incentive Compatible Scheme for Attaining Lindahl Allocations", *Econometrica*, 49, 65-71.

## 7 Appendix

**Proposition 1.** The social choice function  $F(\cdot)$  cannot be implemented in dominant strategies.

**Proof:** By contradiction. To implement  $F(\cdot)$  in dominant strategies there must a strategy set  $S$ , an outcome function  $g : S \rightarrow \{a, b, c\}$ , and strategies  $s_i^j$ , such that  $s_i^j$  is dominant for agent  $i \in \{1, 2, 3\}$  in state  $j \in \{r, y, g\}$  (where  $r$  stands for red,  $y$  for yellow and  $g$  for green).

Since  $g(\cdot)$  implements  $F(\cdot)$  we must have that  $g(s_1^r, s_2^r, s_3^r) = a$ . Since  $a$  is the least favorite outcome for player 2 under state  $r$ , and  $s_2^r$  is dominant for 2 under  $r$ , we must have that  $g(s_1^r, s_2^y, s_3^r) = a$ . Similarly, since  $c$  is the least favorite outcome for player 3 under state  $y$ , and  $s_3^y$  is dominant for 3 under  $y$ , we must have that  $g(s_1^y, s_2^y, s_3^g) = c$ . And also, since  $b$  is the least favorite outcome for player 1 under state  $g$ , and  $s_1^g$  is dominant for 1 under  $g$ , we must have that  $g(s_1^r, s_2^g, s_3^g) = b$ .

Now since  $s_1^r$  is dominant for 1 under  $r$  and  $s_1^y$  is dominant under  $y$  we must have that  $g(s_1^r, s_2^y, s_3^g) \succeq g(s_1^y, s_2^y, s_3^g)$  for player 1 under state  $r$  and  $g(s_1^r, s_2^y, s_3^g) \preceq g(s_1^y, s_2^y, s_3^g)$  for player 1 under state  $y$ . Since we just showed that  $g(s_1^y, s_2^y, s_3^g) = c$ , this implies that  $g(s_1^r, s_2^y, s_3^g) \neq b$ .

Since  $s_2^y$  is dominant for 2 under  $y$  and  $s_2^g$  is dominant under  $g$  we must have that  $g(s_1^r, s_2^y, s_3^g) \succeq g(s_1^r, s_2^g, s_3^g)$  for player 2 under state  $y$  and  $g(s_1^r, s_2^y, s_3^g) \preceq g(s_1^r, s_2^g, s_3^g)$  for player 2 under state  $g$ . Since we just showed that  $g(s_1^r, s_2^g, s_3^g) = b$ , this implies that  $g(s_1^r, s_2^y, s_3^g) \neq a$ .

Since  $s_3^r$  is dominant for 3 under  $r$  and  $s_3^g$  is dominant under  $g$  we must have that  $g(s_1^r, s_2^y, s_3^g) \succeq g(s_1^r, s_2^y, s_3^r)$  for player 3 under state  $g$  and  $g(s_1^r, s_2^y, s_3^g) \preceq g(s_1^r, s_2^y, s_3^r)$  for player 3 under state  $r$ . Since we just showed that  $g(s_1^r, s_2^y, s_3^g) = a$ , this implies that  $g(s_1^r, s_2^y, s_3^g) \neq c$ .

Since  $g(s_1^r, s_2^y, s_3^g) \neq b$ ,  $g(s_1^r, s_2^y, s_3^g) \neq a$ , and  $g(s_1^r, s_2^y, s_3^g) \neq c$ , and there are no other outcomes we reach a contradiction and the result follows.  $\square$

This result is useful to know because it makes apparent the necessity of implementing with a more demanding equilibrium concept. The obvious choice in this case is to implement in Nash equilibrium.

Strategy space: Let  $\Theta = \{\theta_1, \theta_2, \theta_3\}$ ,  $\Lambda = \{a, b, c\}$ ,

**Proposition 2.** The previously described mechanism implements  $F(\cdot)$  in pure strategy Nash equilibrium.

**Proof:** First, notice that a strategy profile in which all agents announce the true state is a Nash equilibrium, as the only agent who can change the outcome in that case is the one who already has her favorite outcome.



Now we show that outcomes which are not desired by the social choice function cannot be the outcome of a pure strategy equilibrium. For this we consider several subcases:

1. Suppose that all agents are announcing untruthfully the same state. In this case there is an agent (agent 1 if the consensus is  $r$ , agent 2 if it is  $y$ , and 3 if it is  $g$ ) which can change the outcome and strictly improve by announcing the true state.
2. Suppose that exactly 2 agents are announcing the same state. One of those agents is not getting her favorite outcome. That agent can change her announcement of the state in such a way that three different states will be announced. She can also choose the integer so that the outcome she announces is selected. If she chooses her favorite outcome she will obtain a strict improvement.
3. Suppose that the outcome is determined by looking at the integers. Then, either of the agents that is not obtaining her favorite outcome can change her announcement of the integer so that the outcome she announces is the one selected. If she also announces her favorite outcome she obtains a strict improvement.

Since this exhausts all cases, the results follows.  $\square$

## INSTRUCTIONS (First treatment)

Thank you for participating in this experiment. In this experiment, there are 10 periods and 3 types of people. The result of one of those periods will determine the money that you will receive in this experiment. We have given you a sheet of paper with spaces to do an announcement in every period. Your identification number and your type are printed on them and will not change during the experiment. In each period, you will be in a group with two other people, so that every group has one person of each type. The other people in your group will not be constant for all ten periods; instead, participants will be re-matched, by identification numbers, with others for each period. While you may be matched with the same person(s) on more than one occasion, you will not know it and at no point will you ever know the identification number or the identity of the other group members in any period.

Your benefits in each period are determined by the combination of the “state of nature” (a color drawn randomly), your “preferences” in that state of nature, and one of the 3 possible “outcomes” that will be the **decreed** by a central processor in each period using the information given by the groups members. The state of nature (red, yellow or green) is obtained randomly at the beginning of the period and is revealed to all the participants. The 3 different types of people have different preferences among the outcomes in each state of nature and consequently different benefits in every case.

Each period you will make an announcement about the state of nature in that period. You can announce any color you wish (it does not have to be the color that was drawn). Your announcement does not change your preferences nor the state of nature, but is part of the information used by the central processor to determine the outcome. The state of nature is the color of a card drawn randomly from a box in which there are 3 yellow cards, 4 green cards, and 5 red cards. The card drawn is shown publicly to everyone in the room. An announcement includes both a color, and outcome and an integer from  $\{1, 2, 3\}$ . The central processor will use the integers and the *announced outcome* to determine the *decreed outcome* when each of the 3 group members announces a different color.

Although these terms are intended to be quite general, here is a specific example:

Consider the state of nature to be the “weather”, the announcement to be a “weather report” and the outcome to be an “activity”. Suppose the weather may be either “hot” (red), “warm” (yellow), or “cold” (green), and that there are 3 possible activities: exercising (a), watching TV (b), and reading (c). Think of the 3 types as three different siblings and the central processor as an absent tutor, who must decide on an activity for her children for the day without knowing the weather, using only the weather reports of her offspring.

If the weather is hot (**red**):

A *type 1* prefers exercise (**a**), next prefers watching TV (**b**), and least prefers reading (**c**).

A *type 2* prefers watching TV (**b**), next prefers reading (**c**), and least prefers exercise (**a**).

A *type 3* prefers reading (**c**), next prefers exercise (**a**), and least prefers watching TV (**b**).

If the weather is warm (**yellow**):

A *type 1* prefers watching TV (**b**), next prefers reading (**c**), and least prefers exercise (**a**).

A *type 2* prefers reading (**c**), next prefers exercise (**a**), and least prefers watching TV (**b**).

A *type 3* prefers exercise (**a**), next prefers watching TV (**b**), and least prefers reading (**c**).

If the weather is cold(**green**):

A *type 1* prefers reading (**c**), next prefers exercise (**a**), and least prefers watching TV (**b**).

A *type 2* prefers exercise (**a**), next prefers watching TV (**b**), and least prefers reading (**c**).

A *type 3* prefers watching TV (**b**), next prefers reading (**c**), and least prefers exercise (**a**).

The following table summarizes this information:

	<i>red</i>	<i>yellow</i>	<i>green</i>
1	$a > b > c$	$b > c > a$	$c > a > b$
2	$b > c > a$	$c > a > b$	$a > b > c$
3	$c > a > b$	$a > b > c$	$b > c > a$

## MONETARY BENEFITS IN PESETAS FOR THE CHOSEN PERIOD

We assume that there is a monetary equivalent for the utility enjoyed by the activities. The 9 statements below describe the money received by the three types of players in each state of nature:

In state **red**,

- a type 1 receives 1500 with outcome  $a$ , 1000 with outcome  $b$ , and 500 with outcome  $c$ .
- a type 2 receives 500 with outcome  $a$ , 1500 with outcome  $b$ , and 1000 with outcome  $c$ .
- a type 3 receives 1000 with outcome  $a$ , 500 with outcome  $b$ , and 1500 with outcome  $c$ .

In state **yellow**:

- a type 1 receives 500 with outcome  $a$ , 1500 with outcome  $b$ , and 1000 with outcome  $c$ .
- a type 2 receives 1000 with outcome  $a$ , 500 with outcome  $b$ , and 1500 with outcome  $c$ .
- a type 3 receives 1500 with outcome  $a$ , 1000 with outcome  $b$ , and 500 with outcome  $c$ .

In state **green**:

- a type 1 receives 1000 with outcome  $a$ , 500 with outcome  $b$ , and 1500 with outcome  $c$ .
- a type 2 receives 1500 with outcome  $a$ , 1000 with outcome  $b$ , and 500 with outcome  $c$ .
- a type 3 receives 500 with outcome  $a$ , 1500 with outcome  $b$ , and 1000 with outcome  $c$ .

## OUTCOME RULES

1. If the 3 group members announce:

**red**, then the outcome is **a**.

**yellow**, then the outcome is **c**.

**green**, then the outcome is **b**.

We can summarize this information in the following way:

**RRR**=*a*, **YYY**=*c*, **GGG**=*a*

The first capital letter denotes the announcement of type 1 (R stands for red, A for yellow and G for green), the second capital letter is the announcement of type 2, the third capital letter is the announcement of type 3, and the lowercase letter after the equal sign denotes the outcome given those announcements.

2. If exactly two group members people announce **red**, the outcome is **a**, unless the group member announcing something different is a *type 1*.

In that case if the *type 1* announces:

**yellow**, the outcome is **b**. **green**, the outcome is **c**.

**RRY**=*a*, **RRG**=*a*, **RYR**=*a*, **RGR**=*a*, **YRR**=*b*, **GRR**=*c*.

3. If exactly two group members people announce **yellow**, the outcome is **c**, unless the group member announcing something different is a *type 2*.

In that case if the *type 2* announces:

**red**, the outcome is **b**. **green**, the outcome is **a**.

**YYR**=*c*, **YYG**=*c*, **RYY**=*c*, **GYG**=*c*, **YRY**=*b*, **YGY**=*a*.

4. If exactly two group members people announce **green**, the outcome is **b**, unless the group member announcing something different is a *type 3*.

In that case if the *type 3* announces:

**red**, the outcome is **c**. **yellow**, the outcome is **a**.

**GYG**=*b*, **GRG**=*b*, **YGG**=*b*, **RGG**=*b*, **GGY**=*a*, **GGR**=*c*.

5. If all 3 members of a group announce different colors, then the central processor adds the three integers selected by the 3 group members. The processor in this case will decree the *announced outcome* (*a*, *b*, or *c*) by one of the group members.

That group member is chosen in the following way:

If the total of the 3 integers chosen is 4 or 7, then the group member is the type 1 person.

= If the total of the 3 integers chosen is 5 or 8, then the group member is the type 2 person.

If the total of the 3 integers chosen is 3, 6, or 9, then the group member is the type 3 person.

**$\text{RYG}(3)=3$ ,  $\text{RYG}(4)=1$ ,  $\text{RYG}(5)=2$ ,  $\text{RYG}(6)=3$ ,  $\text{RYG}(7)=1$ ,  $\text{RYG}(8)=2$ ,  $\text{RYG}(9)=3$ .**

The number in parenthesis to the right of the equal sign is the sum of the announced integers, and the number to the left of the equal sign is the type of the agent whose *announced outcome* will become the *decreed outcome*. The same thing that happens with AVR also happens with ARV, VAR, VRA, RAV, RVA.

(Notice that there are as many combinations which sum to 4 or 7 -exactly 9- as there are for 5 or 8, or even 3, 6 or 9).

## PROCEDURE

When the experiment begins, a color will be randomly drawn and you you will write an announcement in your sheet. The announcement consist of declaring at the same time a state of the world (red, yellow or green), an integer in  $\{1, 2, 3\}$ , and an outcome in  $\{a, b, c\}$ .

The announcement sheets will then be collected and the announcements will be processed to determine the outcome, either  $a$ ,  $b$ , or  $c$ .

The experimenter will then compute your benefits for the period and your announcement sheet will be returned to you with these indicated. You will only be informed of your payoffs. You will not be informed of the announcements or payoffs of other group members.

Next we will proceed to the following period. At the end of 10 periods, the experiment will end. Each person will receive a show-up fee and the the benefits obtained in the period selected to be the payment period. Each person will be paid individually and privately.

The payment period will be chosen at random at the end of the experiment. We will have cards numbered from 0 to 10. A student will select one of these cards at random and the number of the card selected will determine the payment period.

## EXERCISES

To ensure that people understand how the mechanism works, we will do some exercises.

1. Suppose that the monitor draws a red card and all group members announce *red*. What is the outcome? What is the state of nature? What are the payoffs for the type 1 person? the type 2 person? the type 3 person?=  
=
2. Suppose that the monitor draws a red card and types 2 and 3 announce *red*, while type 1 announces *green*. What is the outcome? What is the state of nature? What are the payoffs for the type 1 person? the type 2 person? the type 3 person?
3. Suppose that the monitor draws a green card and types 1 and 3 announce *green*, while type 2 announces *red*. What is the outcome? What is the state of nature? What are the payoffs for the type 1 person? the type 2 person? the type 3 person?
4. Suppose that the monitor draws a green card and all group members announce *yellow*? What is the outcome? What are the payoffs for the type 1 person? the type 2 person? the type 3 person?
5. Suppose that the monitor draws a green card and types 1 and 3 announce *yellow*, while type 2 announces *green*. What is the outcome? What are the payoffs for the type 1 person? the type 2 person? the type 3 person?
6. Suppose the monitor draws a yellow card, types 1, 2, and 3 announce (respectively) *red*, *yellow*, and *green*, the integers 1, 2, and 3, and the outcomes *a*, *b*, and *c*. What is the outcome? What are the payoffs for the type 1 person?=  
=20 the type 2 person?  
the type 3 person?
7. Suppose the monitor draws a yellow card, types 1, 2, and 3 announce (respectively) *red*, *yellow*, and *green*, the integers 1, 2, and 2, and the outcomes *c*, *b*, and *c*. What is the outcome? What are the payoffs for the type 1 person?=  
=20 the type 2 person?  
the type 3 person?

Once the experiment begins, all communication between participants is strictly forbidden. Please ask questions before we begin. Are there any questions?

The following tables may be of help in summarizing the information about payoffs.

Announcements	True state		
	R	Y	G
RRR	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
RRY	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
RRG	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
RYR	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
RGR	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
YRR	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
GRR	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
YYY	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
YYR	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
YYG	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
GYG	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
RYY	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
YGY	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
YRY	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
GGG	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
YGG	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
RGG	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
GYG	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
GRG	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
GGY	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
GGR	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000



Payments when the three group members announce different states.

Sum of integers	Selected type and announces outcome	True state		
		R	Y	G
3	3, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
3	3, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
3	3, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
4	1, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
4	1, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
4	1, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
5	2, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
5	2, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
5	2, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
6	3, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
6	3, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
6	3, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
7	1, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
7	1, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
7	1, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
8	2, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
8	2, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
8	2, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
9	3, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
9	3, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
9	3, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000

## INSTRUCTIONS (Second treatment)

Thank you for participating in this experiment. In this experiment, there are 10 periods and 3 types of people. The result of one of those periods will determine the money that you will receive in this experiment. We have given you a sheet of paper with spaces to do an announcement in every period. Your identification number and your type are printed on them and will not change during the experiment. In each period, you will be in a group with two other people, so that every group has one person of each type. The other people in your group will not be constant for all ten periods; instead, participants will be re-matched, by identification numbers, with others for each period. While you may be matched with the same person(s) on more than one occasion, you will not know it and at no point will you ever know the identification number or the identity of the other group members in any period.

Your benefits in each period are determined by the combination of the “state of nature” (a color drawn randomly), your “preferences” in that state of nature, and one of the 3 possible “outcomes” that will be the **decreed** by a central processor in each period using the information given by the groups members. The state of nature (red, yellow or green) is obtained randomly at the beginning of the period and is revealed to all the participants. The 3 different types of people have different preferences among the outcomes in each state of nature and consequently different benefits in every case.

Each period you will make an announcement about the state of nature in that period. You can announce any color you wish (it does not have to be the color that was drawn). Your announcement does not change your preferences nor the state of nature, but is part of the information used by the central processor to determine the outcome. The state of nature is the color of a card drawn randomly from a box in which there are 3 yellow cards, 4 green cards, and 5 red cards. The card drawn is shown publicly to everyone in the room. An announcement includes both a color, and outcome and an integer from  $\{1, 2, 3\}$ . The central processor will use the integers and the *announced outcome* to determine the *decreed outcome* when each of the 3 group members announces a different color.

Although these terms are intended to be quite general, here is a specific example:

Consider the state of nature to be the “weather”, the announcement to be a ”weather report” and the outcome to be an “activity”. Suppose the weather may be either “hot” (red), “warm” (yellow), or “cold” (green), and that there are 3 possible activities: exercising (a), watching TV (b), and reading (c). Think of the 3 types as three different siblings and the central processor as an absent tutor, who must decide on an activity for her children for the day without knowing the weather, using only the weather reports of her offspring.

If the weather is hot (**red**):

A *type 1* prefers exercise (**a**), next prefers watching TV (**b**), and least prefers reading (**c**).

A *type 2* prefers watching TV (**b**), next prefers reading (**c**), and least prefers exercise (**a**).

A *type 3* prefers reading (**c**), next prefers exercise (**a**), and least prefers watching TV (**b**).

If the weather is warm (**yellow**):

A *type 1* prefers watching TV (**b**), next prefers reading (**c**), and least prefers exercise (**a**).

A *type 2* prefers reading (**c**), next prefers exercise (**a**), and least prefers watching TV (**b**).

A *type 3* prefers exercise (**a**), next prefers watching TV (**b**), and least prefers reading (**c**).

If the weather is cold(**green**):

A *type 1* prefers reading (**c**), next prefers exercise (**a**), and least prefers watching TV (**b**).

A *type 2* prefers exercise (**a**), next prefers watching TV (**b**), and least prefers reading (**c**).

A *type 3* prefers watching TV (**b**), next prefers reading (**c**), and least prefers exercise (**a**).

The following table summarizes this information:

	<i>red</i>	<i>yellow</i>	<i>green</i>
1	$a > b > c$	$b > c > a$	$c > a > b$
2	$b > c > a$	$c > a > b$	$a > b > c$
3	$c > a > b$	$a > b > c$	$b > c > a$

## MONETARY BENEFITS IN PESETAS FOR THE CHOSEN PERIOD

We assume that there is a monetary equivalent for the utility enjoyed by the activities. The 9 statements below describe the money received by the three types of players in each state of nature:

In state **red**,

a type 1 receives 1500 with outcome  $a$ , 1000 with outcome  $b$ , and 500 with outcome  $c$ .  
a type 2 receives 500 with outcome  $a$ , 1500 with outcome  $b$ , and 1000 with outcome  $c$ .  
a type 3 receives 1000 with outcome  $a$ , 500 with outcome  $b$ , and 1500 with outcome  $c$ .

In state **yellow**:

a type 1 receives 500 with outcome  $a$ , 1500 with outcome  $b$ , and 1000 with outcome  $c$ .  
a type 2 receives 1000 with outcome  $a$ , 500 with outcome  $b$ , and 1500 with outcome  $c$ .  
a type 3 receives 1500 with outcome  $a$ , 1000 with outcome  $b$ , and 500 with outcome  $c$ .

In state **green**:

a type 1 receives 1000 with outcome  $a$ , 500 with outcome  $b$ , and 1500 with outcome  $c$ .  
a type 2 receives 1500 with outcome  $a$ , 1000 with outcome  $b$ , and 500 with outcome  $c$ .  
a type 3 receives 500 with outcome  $a$ , 1500 with outcome  $b$ , and 1000 with outcome  $c$ .

The gains described here will be modified in some cases, as described in points 2, 3 and 4 of page 3.

## OUTCOME RULES

1. If the 3 group members announce:

**red**, then the outcome is **a**.

**yellow**, then the outcome is **c**.

**green**, then the outcome is **b**.

We can summarize this information in the following way:

**RRR**=*a*, **YYY**=*c*, **GGG**=*a*

The first capital letter denotes the announcement of type 1 (R stands for red, A for yellow and G for green), the second capital letter is the announcement of type 2, the third capital letter is the announcement of type 3, and the lowercase letter after the equal sign denotes the outcome given those announcements.

2. If exactly two group members people announce **red**, the outcome is **a**, unless the group member announcing something different is a *type* 1.

In that case if the *type* 1 announces:

**yellow**, the outcome is **b**. **green**, the outcome is **c**.

**RRY**=*a*, **RRG**=*a*, **RYR**=*a*, **RGR**=*a*, **YRR**=*b*, **GRR**=*c*.

If the announcement is **RRY** or **RRG**, then the group member of *type* 3 will receive **200 pesetas less** than the quantity shown in page 3. If the announcement is **RYR** or **RGR**, then the group member of *type* 2 will receive **100 pesetas less** than the quantity shown in page 3. This is illustrated in the table of page 7.

3. If exactly two group members people announce **yellow**, the outcome is **c**, unless the group member announcing something different is a *type* 2.

In that case if the *type* 2 announces:

**red**, the outcome is **b**. **green**, the outcome is **a**.

**YYR**=*c*, **YYG**=*c*, **RYY**=*c*, **GYY**=*c*, **YRY**=*b*, **YGY**=*a*.

If the announcement is **RYY** or **GYY**, then the group member of *type* 1 will receive **200 pesetas less** than the quantity shown in page 3. If the announcement is **YYR** or **YYG**, then the group member of *type* 2 will receive **100 pesetas less** than the quantity shown in page 3. This is illustrated in the table of page 7.

4. If exactly two group members people announce **green**, the outcome is **b**, unless the group member announcing something different is a *type* 3.

In that case if the *type* 3 announces:

**red**, the outcome is **c**. **yellow**, the outcome is **a**.

**GYG**=*b*, **GRG**=*b*, **YGG**=*b*, **RGG**=*b*, **GGY**=*a*, **GGR**=*c*.

If the announcement is **GYG** or **GRG**, then the group member of *type 2* will receive **200 pesetas less** than the quantity shown in page 3. If the announcement is **YGG** or **RGG**, then the group member of *type 1* will receive **100 pesetas less** than the quantity shown in page 3. This is illustrated in the table of page 7.

5. If all 3 members of a group announce different colors, then the central processor adds the three integers selected by the 3 group members. The processor in this case will decree the *announced outcome* ( $a$ ,  $b$ , or  $c$ ) by one of the group members.

That group member is chosen in the following way:

If the total of the 3 integers chosen is 4 or 7, then the group member is the type 1 person.

= If the total of the 3 integers chosen is 5 or 8, then the group member is the type 2 person.

If the total of the 3 integers chosen is 3, 6, or 9, then the group member is the type 3 person.

**RYG(3)=3, RYG(4)=1, RYG(5)=2, RYG(6)=3, RYG(7)=1, RYG(8)=2, RYG(9)=3.**

The number in parenthesis to the right of the equal sign is the sum of the announced integers, and the number to the left of the equal sign is the type of the agent whose *announced outcome* will become the *decreed outcome*. The same thing that happens with AVR also happens with ARV, VAR, VRA, RAV, RVA.

(Notice that there are as many combinations which sum to 4 or 7 -exactly 9- as there are for 5 or 8, or even 3, 6 or 9).

## PROCEDURE

When the experiment begins, a color will be randomly drawn and you you will write an announcement in your sheet. The announcement consist of declaring at the same time a state of the world (red, yellow or green), an integer in  $\{1, 2, 3\}$ , and an outcome in  $\{a, b, c\}$ .

The announcement sheets will then be collected and the announcements will be processed to determine the outcome, either  $a$ ,  $b$ , or  $c$ .

The experimenter will then compute your benefits for the period and your announcement sheet will be returned to you with these indicated. You will only be informed of your payoffs. You will not be informed of the announcements or payoffs of other group members.

Next we will proceed to the following period. At the end of 10 periods, the experiment will end. Each person will receive a show-up fee and the the benefits obtained in the period selected to be the payment period. Each person will be paid individually and privately.

The payment period will be chosen at random at the end of the experiment. We will have cards numbered from 0 to 10. A student will select one of these cards at random and the number of the card selected will determine the payment period.

## EXERCISES

To ensure that people understand how the mechanism works, we will do some exercises.

1. Suppose that the monitor draws a red card and all group members announce *red*. What is the outcome? What is the state of nature? What are the payoffs for the type 1 person? the type 2 person? the type 3 person?=  
=
2. Suppose that the monitor draws a red card and types 2 and 3 announce *red*, while type 1 announces *green*. What is the outcome? What is the state of nature? What are the payoffs for the type 1 person? the type 2 person? the type 3 person?
3. Suppose that the monitor draws a green card and types 1 and 3 announce *green*, while type 2 announces *red*. What is the outcome? What is the state of nature? What are the payoffs for the type 1 person? the type 2 person? the type 3 person?
4. Suppose that the monitor draws a green card and all group members announce *yellow*? What is the outcome? What are the payoffs for the type 1 person? the type 2 person? the type 3 person?
5. Suppose that the monitor draws a green card and types 1 and 3 announce *yellow*, while type 2 announces *green*. What is the outcome? What are the payoffs for the type 1 person? the type 2 person? the type 3 person?
6. Suppose the monitor draws a yellow card, types 1, 2, and 3 announce (respectively) *red*, *yellow*, and *green*, the integers 1, 2, and 3, and the outcomes  $a$ ,  $b$ , and  $c$ . What is the outcome? What are the payoffs for the type 1 person?= $20$  the type 2 person? the type 3 person?
7. Suppose the monitor draws a yellow card, types 1, 2, and 3 announce (respectively) *red*, *yellow*, and *green*, the integers 1, 2, and 2, and the outcomes  $c$ ,  $b$ , and  $c$ . What is the outcome? What are the payoffs for the type 1 person?= $20$  the type 2 person? the type 3 person?

Once the experiment begins, all communication between participants is strictly forbidden. Please ask questions before we begin. Are there any questions?

The following tables may be of help in summarizing the information about payoffs.

Announcements	True state		
	R	Y	G
RRR	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
RRY	1500, 500, 800	500, 1000, 1300	1000, 1500, 300
RRG	1500, 500, 800	500, 1000, 1300	1000, 1500, 300
RYR	1500, 400, 1000	500, 900, 1500	1000, 1400, 500
RGR	1500, 400, 1000	500, 900, 1500	1000, 1400, 500
YRR	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
GRR	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
YYY	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
YYR	500, 1000, 1400	1000, 1500, 400	1500, 500, 900
YYG	500, 1000, 1400	1000, 1500, 400	1500, 500, 900
GYR	300, 1000, 1500	800, 1500, 500	1300, 500, 1000
RYY	300, 1000, 1500	800, 1500, 500	1300, 500, 1000
YGY	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
YRY	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
GGG	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
YGG	900, 1500, 500	1400, 500, 1000	400, 1000, 1500
RGG	900, 1500, 500	1400, 500, 1000	400, 1000, 1500
GYG	1000, 1300, 500	1500, 300, 1000	500, 800, 1500
GRG	1000, 1300, 500	1500, 300, 1000	500, 800, 1500
GGY	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
GGR	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000



Payments when the three group members announce different states.

Sum of integers	Selected type and announces outcome	True state		
		R	Y	G
3	3, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
3	3, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
3	3, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
4	1, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
4	1, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
4	1, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
5	2, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
5	2, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
5	2, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
6	3, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
6	3, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
6	3, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
7	1, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
7	1, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
7	1, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
8	2, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
8	2, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
8	2, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000
9	3, <i>a</i>	1500, 500, 1000	500, 1000, 1500	1000, 1500, 500
9	3, <i>b</i>	1000, 1500, 500	1500, 500, 1000	500, 1000, 1500
9	3, <i>c</i>	500, 1000, 1500	1000, 1500, 500	1500, 500, 1000

## RESULTS

Baseline Treatment

Session 1

Period	True State	Group										Number of Socially Preferred Outcomes
		1		2		3		4		5		
		Ann. States	Out.	Ann. States	Out.	Ann. States	Out.	Ann. States	Out.	Ann. States	Out.	
1	Y	RYR	a	YRR	b	YYY	c	GYY	c	YYY	c	3
2	Y	YYG	c	RYG	a	YYY	c	YYR	c	RYG	b	3
3	G	GRR	c	GGG	b	GRG	b	YRG	a	YRG	c	2
4	G	GGG	b	GGG	b	YRR	b	YGG	b	YRG	c	4
5	R	RGY	b	RGR	a	RGG	b	RYR	a	GGY	a	3
6	G	GRR	c	RGG	b	RGG	b	YGG	b	YGG	b	4
7	Y	YYY	c	GYR	b	YYR	c	YYY	c	GYY	c	4
8	R	RYR	a	RYR	a	RYR	a	RGY	b	RGR	a	4
9	G	GGG	b	YGG	b	RGG	b	YGG	b	YGG	b	5
10	R	RGR	a	RYY	c	RGR	a	RYG	c	RGR	a	3

35

Key: R=red, Y=yellow, G=green. Each triplet of announced states is in the order type1, type2, type 3 for that group.

Baseline Treatment

Session 2

Period	True State	Group										Number of Socially Preferred Outcomes
		1		2		3		4		5		
		Ann. States	Out.	Ann. States	Out.	Ann. States	Out.	Ann. States	Out.	Ann. States	Out.	
1	R	RRG	a	RRR	a	RRR	a	RGR	a	RRR	a	5
2	R	RRY	a	RGY	a	YGY	a	RGY	b	RGR	a	4
3	G	YGG	b	GGG	b	GRG	b	RRG	a	YRG	c	3
4	R	RGR	a	RYR	a	RRY	a	RGY	b	RYY	c	3
5	Y	YYR	c	YYR	c	YYR	c	YYG	c	GGR	c	5
6	R	RGR	a	RYR	a	RGY	a	RYR	a	RYR	a	5
7	R	RGY	b	RYY	c	RYY	c	RGR	a	RYR	a	2
8	G	GRG	b	RRG	a	YRG	a	RRG	a	YGG	b	2
9	Y	GYG	b	RYR	a	YYR	c	YYR	c	YYG	c	3
10	R	RYY	c	RGY	c	RGY	b	RGR	a	RYY	c	1
												33

Key: R=red, Y=yellow, G=green. Each triplet of announced states is in the order type1, type2, type 3 for that group.

Fine Treatment

Session 1

Period	True State	Group										Number of Socially Preferred Outcomes
		1		2		3		4		5		
		Ann. States	Out.	Ann. States	Out.	Ann. States	Out.	Ann. States	Out.	Ann. States	Out.	
1	R	RRR	a	RYY	c	RGR	a	RGY	a	RGR	a	4
2	R	RYR	a	YRG	a	RRY	a	RYR	a	RGR	a	5
3	G	YGG	b	YGG	b	GRR	c	GRG	b	YGG	b	4
4	Y	YYR	c	YYG	c	YYG	c	YGY	a	GYR	a	3
5	R	RYY	c	YYR	c	RYR	a	RRR	a	RRY	a	3
6	Y	YGY	a	GYG	b	YYY	c	YYR	c	GYY	c	3
7	Y	YYG	c	GYR	b	YYR	c	YYR	c	GYY	c	4
8	G	RGG	b	RGG	b	GGG	b	YGG	b	GGG	b	5
9	G	GGG	b	GGG	b	GGG	b	RRG	a	RGG	b	4
10	Y	GYY	c	GYY	c	YYG	c	GYR	a	YYG	c	4
												39

Key: R=red, Y=yellow, G=green. Each triplet of announced states is in the order type1, type2, type 3 for that group.

Fine Treatment

Session 2

Period	True State	Group										Number of Socially Preferred Outcomes
		1		2		3		4		5		
		Ann. States	Out.	Ann. States	Out.	Ann. States	Out.	Ann. States	Out.	Ann. States	Out.	
1	Y	YYY	c	GYG	c	YYR	c	GYG	c	YYY	c	5
2	Y	YYG	c	GGY	a	GYG	b	YGY	a	GYG	c	2
3	R	RRR	a	RYY	a	RYY	a	RYY	a	RRR	a	5
4	R	RGR	a	RRR	a	RRR	a	RRR	a	RRR	a	5
5	G	GRG	b	RGG	b	YGG	b	YRG	b	YGG	b	5
6	Y	YYG	c	YYR	c	GYG	c	GYG	c	YGY	a	4
7	R	RGR	a	RRY	a	RYY	a	RYY	c	RRR	a	4
8	G	GRG	b	GRG	b	RRG	a	RGG	b	GGG	b	4
9	G	RRG	a	GGG	b	RGG	b	GGR	c	GGG	b	3
10	Y	YYR	c	YGR	c	YYR	c	GYG	b	RYY	c	4
												41

Key: R=red, Y=yellow, G=green. Each triplet of announced states is in the order type1, type2, type 3 for that group.

**COLORS REPORTED BY 2ND-FAVORED AND LEAST FAVORED PLAYER  
UNDER  $F(\cdot)$**

**No Fine**

True State = R

Session/Period	# R by 3	# Y by 3	# Y by 2	# G by 2
1/5	2	2	1	4
1/8	4	1	3	2
1/10	3	1	2	3
2/1	4	0	0	1
2/2	1	4	0	4
2/4	2	3	2	2
2/6	4	1	3	2
2/7	2	3	3	2
2/10	1	4	2	3
Totals	23	19	16	23

True State = Y

Session/Period	# Y by 1	# G by 1	# G by 3	# R by 3
1/1	3	1	0	2
1/2	3	1	2	2
1/7	3	2	0	2
2/5	4	1	1	4
2/9	3	1	2	3
Totals	16	6	5	13

True State = G

Session/Period	# G by 2	# R by 2	# R by 1	# Y by 1
1/3	0	5	0	2
1/4	3	2	0	3
1/6	4	1	2	2
1/9	5	0	1	3
2/5	2	3	1	2
2/9	1	4	2	2
Totals	15	15	6	14

## Fine

True State = R

Session/Period	# R by 3	# Y by 3	# Y by 2	# G by 2
3/1	3	2	1	3
3/2	3	1	2	1
3/5	3	2	3	0
4/3	5	0	3	0
4/4	5	0	0	1
4/7	3	1	2	1
Totals	22	6	11	6

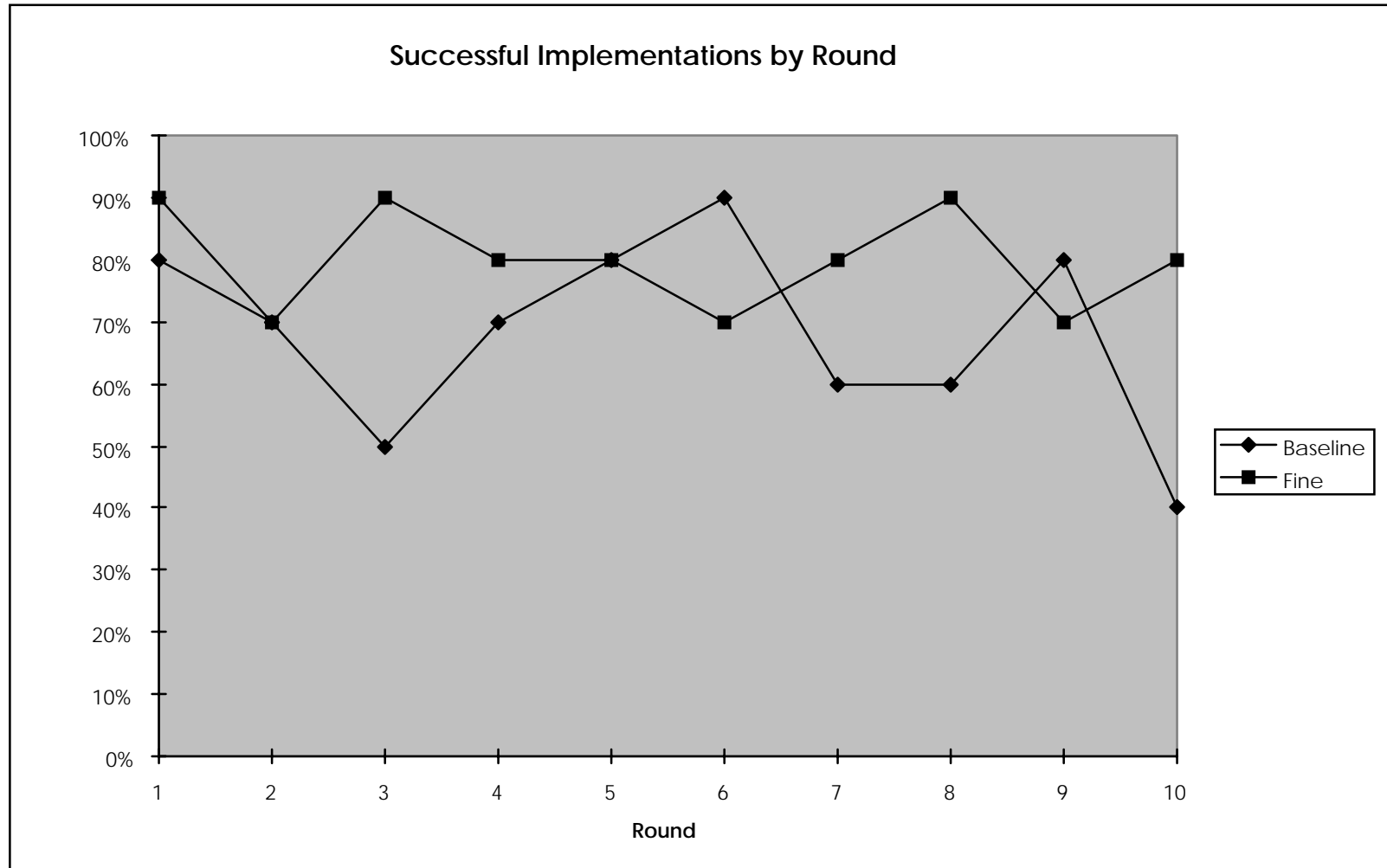
True State = Y

Session/Period	# Y by 1	# G by 1	# G by 3	# R by 3
3/4	4	1	2	2
3/6	3	2	1	1
3/7	3	2	1	3
3/10	2	3	2	1
4/1	3	2	0	1
4/2	2	3	2	0
4/6	3	2	1	1
4/10	3	1	1	3
Totals	23	16	12	10

True State = G

Session/Period	# G by 2	# R by 2	# R by 1	# Y by 1
3/3	3	2	0	3
3/8	5	0	2	1
3/9	4	1	2	0
4/5	3	2	1	3
4/8	2	3	2	0
4/9	4	1	2	0
Totals	21	9	9	7

**FIGURE 1**





**FIGURE 2**

