Experiments on Emergent Leadership, Lying Aversion, and Reciprocal Altruism: The Importance of Context[*]

Pablo Hernandez, UC Berkeley-Haas

September 19, 2012

Abstract

From an experimental design, we find that circumstance determines who takes the lead through communication. When the context varies exogenously, from containing mild to severe strategic conflict, initiative is taken by anyone in the former case, but by special individuals in the latter case. These special leaders dislike free-riding and lying. We show that initiative leads to cooperation even in games of severe strategic conflict such as the Prisoners' Dilemma.

1 Introduction

The Arab spring, the Occupy movement, the Indignants movement in Spain and the protests in Moscow following Putin's re-election all relied on leaderless groups coalescing and demonstrating out on the streets. While these groups lacked formal leaders, informal leaders quickly emerged. These were the individuals using communication tools like Twitter to organize individuals in the protests. Who are these individuals? What characteristics separate them from others? How does the context of the government's reaction to the demonstrators affect the success of organizing as well as the identity of the organizers? To investigate this, we report the results of controlled laboratory experiments on spontaneous leadership—the emergence of leaders from leaderless groups. While many characteristics have been shown to be related to leadership, we highlight two novel features, lying aversion and reciprocal altruism that determine leadership in context.

To be precise, consider a simple setting where two individuals must organize to overthrow the regime. If both choose to demonstrate, the government falls, and this is the best outcome from a

[*] Preliminary draft, please do not cite.

societal perspective. There is, however, a risk in demonstrating. If only one demonstrator turns up, then the government represses that individual and he or she suffers the worst possible outcome. A demonstrator can avoid this risk by staying home. When everyone stays home, the regime remains in place and both individuals obtain lower payoffs than had they both demonstrated. The key variation in context concerns what happens to "loyalists" who stay home in the face of demonstrations from others. One possibility is that they receive no additional reward. The regime remains in place and they are as well off as under the circumstances where no one protested. Alternatively, the regime might choose to reward loyalists in this situation. Here, we imagine that the reward is so large that the loyalist is better off when the regime stays in place in the face of demonstrations than when the regime falls.

These two situations may be readily recognized as a stag hunt game (a coordination game in which defection is risk dominant) and a prisoner's dilemma game. Leadership consists of initiating messages to organize demonstrations. In the stag hunt, it is an equilibrium for everyone to demonstrate though there is obviously some risk involved. Here, the role of a leader is to create the trust necessary for everyone to take to the streets. In contrast, in the prisoner's dilemma demonstrating by everyone is a dominated strategy. Here, a leader must convince others to override their self-interest, but there is temptation present for a successful leader. By rallying the masses to the streets and then staying home, the leader can benefit from the reward for loyalty.[1]

So how does leadership in these two situations relate to lying aversion and reciprocal altruism? Lying aversion is the cost of breaking one's word. Since honoring one's word is optimal in stag hunt, there is no tension here. Leadership decisions do not hinge on this trait. For the prisoner's dilemma, the situation is quite different. The temptation of receiving a reward for loyalty (and of avoiding the possibility of severe punishment) provides an incentive for a leader to break his or her word. Thus, potential leaders who anticipate that they will succumb to this temptation may choose to eschew leadership altogether. On the other hand, once committed to leadership, lying aversion provides commitment to follow through and this may make a leader more persuasive in rallying others to demonstrate. In short, the effect of lying aversion on leadership is context dependent.

Reciprocal altruism is the idea that a leader gets intrinsic value when others join him or her in demonstrating. In the stag hunt, this effect merely reinforces the extrinsic incentives already present

[1] Evidence suggests that the Occupy movement is closer to our stag hunt stylization and the Tunisian and Egyptian uprising to our prisoner's dilemma. The allegedly non-violent character of the Occupy movement made it less risky for those who decided to follow. The Tunisian and Egyptian uprisings were explicitly oriented to overthrow their corresponding regimes –followers would have faced a much higher cost the protests had not been successful.

and so again it is of no consequence. In contrast, reciprocal altruism undermines the temptation to stay home when others are out demonstrating in the prisoner's dilemma. Thus, it undercuts the incentive effects of the rewards to loyalists. These additional gains drive individuals with high intrinsic benefits to leadership roles.

Moreover, there is potentially an interaction effect. An individual who is both lying averse and reciprocally altruistic is most likely to take a leadership role since the two effects reinforce one another—there is no temptation to back out of the demonstration and ample reward for rallying others. A lying averse but not reciprocally altruistic individual will shy away from leadership since there is no particular benefit to rallying others and a substantial cost should the individual succumb to temptation and stay at home. An individual who is neither lying averse nor reciprocally altruistic is more likely to take a leadership role. Here, the individual is, in effect, acting as an agent of the state and seeking to "out" the state's enemies in exchange for the rewards from loyalty. The point is simply that context is crucial as to whether these characteristics are activated in determining the identity of leaders.

In this paper, we perform laboratory experiments to study the relationship between leadership, lying aversion, and reciprocal altruism. We perform a battery of tests to measure these characteristics as well as a host of others that have been found to be indicative of leadership. The heart of the study is to place subjects in leaderless groups where they have an opportunity to communicate. We observe the identity of the person initiating a plan for mass demonstrations, observe whether the plan is agreed to, and then study subsequent actions. The key treatment variation is the context in which leadership takes place, i.e., the regime's reward scheme for loyalists. In our baseline treatment, the game is a stag hunt—the regime offers no additional rewards to loyalists. We compare this to a treatment where the regime lavishly rewards loyalists in the face of demonstrations, i.e. a prisoner's dilemma situation.

In the baseline treatment, the new leadership characteristics we study have no bearing on the identity of the person initiating the plan to demonstrate. Leadership is ubiquitous in this setting and the regime falls the overwhelming majority of the time. In the prisoner's dilemma treatment, these characteristics strongly predict leadership as suggested by the thought experiment. An individual who is lying averse and reciprocally altruistic is most likely to emerge as a leader. A lying averse but not reciprocally altruistic individual is least likely. We do not, however, see much evidence of agents of the state, individuals who lead and then subsequently defect in the game. Compared to the stag hunt situation, the regime falls less often when it rewards loyalists, but still fails 23% of the time, considerably above what one would expect were individuals mainly motivated by money. This offers indirect support for the importance of the new effects we identify.

Throughout the paper we use this analogy about groups coalescing and demonstrating out on the streets, althought the games we study could represent other situations as well. For instance, the documented delevel of organizations from hierarchies to business-to-business teams might give rise to different types of leaders. To take the initiative to set the time and place for a meeting might be associated with different leadership traits than to exhort others to perform an unverifiable action. In the former case there is no individual incentive to go to another place at the time of the meeting, whereas in the latter, individuals (leaders inclusive) may succumb to the temptation to free-ride on others' effort, just as individuals in the uprising states may fail to resist the temptation to staying at home.

The paper is divided as follows. In Section 2 we provide a literature review, in Section 3 we describe the experimental design and in Section 4 the hypotheses. In Section 5 we show the results and in Section 6 we conclude.

2 Literature

The literature on leadership is vast, with great many different definitions and approaches to the phenomenon. It seems however, there is consensus in one idea: Leadership is the natural outcome from strategic interaction in which individuals seek to coordinate towards a common objective (Lewis 1974, Boehm 1999, Van Vugt 2006). This natural outcome manifests itself in a subset of individuals' particular behavior that aims (effectively or not) at achieving shared goals. The behavioral theories of leadership seek for sound classifications of such behavior. The seminal behavioral theories of leadership that have influenced subsequent work can be grouped in two broad categories (DeRue 2011): transformational-transactional leadership, and leadership as consideration and initiating structure. In one hand, the transactional-transformational approach, introduced by Burns (1978), analyzes leaders' actions oriented to manage rewards and punishments to achieve collective goals (the transactional approach), and actions oriented to produce a significant change in people's lifes (the transformational approach). These approaches focus on behavior of already established leaders. In the other hand, the consideration and initiating structure approach, born from the studies on leadership emergence in leaderless groups during the 50s (Bass 1949, 1954; Hemphill 1950; Stogdill 1963), leaders establish plans to coordinate group's actions towards shared goals (initiating structure) taking into account individuals' needs (consideration).

This behavioral approach to leadership was born to complement the more traditional trait theories of leadership that started back in the 19th century (Galton and Eysenck, 1869). The rich literature on leaders' traits has documented several distinctive features of leaders. In a nutshell, emotional traits such as (the feeling of) power, ambition, extroversion (Bass 1990, Judge et al. 2002); or abilities and skills such as energy, intelligence, verbal fluency, confidence and independence (Bass 1990, Avolio, Sosik, Jung and Berson 2003) have been found to be related to leadership emergence and effectiveness.

Perhaps surprisingly, the behavioral perspective on leadership has developed mostly in parallel with the traits perspective, even though scholars now agree that both approaches are essentially related (DeRue 2011). In effect, researchers often lament the lack of integration between the behavioral and trait based approaches (Avolio 2007). The economic view of leadership integrates these two tightly related perspectives because it posits "traits" as the primitives driving behavior, given incentives. This more integrative view is the one we embrace in this study.

Initiative may be driven by particular traits depending on the degree of strategic conflict that circumstance poses to individuals. In coordination games, leadership was stressed first in Kreps (1990). Its main idea posits leaders as the ones who can coordinate in the presence of multiple equilibria. Although interesting, little subsequent work has been done (for a clear discussion see Hermalin 2000) in analyzing leaders as coordinators in a way consistent with social psychology findings (Van Vugt 2006, Foss 2001). In social dilemmas, the most common form of leadership found in the literature is "leading-by-example." Hermalin (1998) shows that individuals endowed with more information about the value of a public good, could lead-by-example or lead-by-sacrifice in order to signal this information. Meidinger and Villeval (2002) experimentally tests Hermalin's theory and finds that although signaling plays a role, reciprocity (mimicking leaders' contribution) provides a better rationale for the data, in both leading-by-example and leading-by-sacrifice. Along the same lines Güth, Levati, Sutter and van der Heijden (2006), Moxnes and van der Heijden (2007), Gachter and Renner (2006) study sequential contributions to a public good. All of them find that letting one member to contribute first raises contributions, mainly because of the large contributions of these leaders. These studies focus on exogenously imposed roles of leaders and followers. There has been research on endogenous leading-by-example as well, mainly on charitable giving (as a version of a public goods game). Potters, Sefton and Vesterlund (2005) for instance, keep the asymmetry of information assumption, but analyze endogenous sequence of contributions. They find that endogenous sequential contribution also improves outcomes. These studies however, do not put emphasis on leaders (and non-leaders) traits.

To my knowledge, there are three studies in which traits and endogenous initiative (by example) are put together in a public goods game. The first one is Bruttel and Fischbacher (2010). By means of a novel experimental design the authors identify costly endogenous leadership behavior and elicit individuals' traits. Males, above-average cognitive skills, generosity and strong preferences for efficiency and equality are related to individuals who take the initiative. Perhaps surprisingly, they do not find an association between personality traits and risk aversion on leadership behavior. The second one is by Rivas and Sutter (2011). In a concise study, they find voluntary leadership is more efficient than exogenously imposed leadership. The third study is Arbak and Villeval (2011). It consists of a public good experiment similar to Rivas and Sutter (2011), that also measures generosity, gender and personality traits. The main treatment is to reveal group members' attributes (gender and generosity) when contribution is voluntary. The authors elicited generosity by asking individuals to give a portion of their show-up fee to a charity, and personality traits through the Big 5 personality test (John 1990). They also find that leading voluntary contributions yield to more efficient outcomes, especially in groups with a high number of generous individuals. Males are more likely to be leaders, and females contribute more on average.

Overall, these papers find that initiative taking (by example) improves outcomes and that other-regarding concerns influence the decision to lead. They however, do not explore the role of strategic conflict and they do not allow for communication. We believe that strategic conflict is fundamental in leadership emergence and that communication is perhaps the most pervasive mechanism of exhortation. Ours is, to my knowledge, the first study addressing the former issue. Regarding the latter however, research is extensive, although no particular emphasis has been placed on endogenous initiative. From a theoretical perspective, Crawford and Sobel (1982) may be regarded as the seminal paper that shows information can be transmitted through cheap-talk. Their result hinges on the partial alignment of interests among two parties. More generally, pre-play communication can be considered as a "device" by which individuals can coordinate in some [Nash] equilibrium (Forges 1986, Barany 1992) of the normal form game. These predictions provide a rationale for equilibrium selection in games with multiple equilibria, such as our stag hunt game. The theoretical literature on pre-play communication however, fails to explain cooperation in social dilemmas, when only pecuniary motives are considered.

Motives can also have a non-pecuaniary component. Kreps, Milgrom, Roberts and Wilson (1982) provide a theoretical explanation for cooperation in a finitely repeated prisoner's dilemma based on the existence of individuals who realise a non-monetary gain when mutual cooperation occurs. Andreoni and Miller (1993) shows that Kreps et al (1982) non-monetary concern or "reciprocal-altruism" explains the data from an experimental design. More recently, Bolton and Ockenfels (2000) provides a model based on other regarding preferences (that imbeds a taste for reciprocity and for fairness) to explain the positive correlation between wage offers and subsequent effort found in the literature (among other non-standard economic behavior). In sum, cooperation in social dilemmas have been found experimentally even without communication (Fehr, 1993; Andreoni and Miller 1993; Cooper, 1996).

When communication is allowed, a robust finding in experimental studies of social dilemmas (for surveys see Ledyard 1994, Sally 1995 and Crawford, 1998) and of coordination games (Cooper 1992) is that pre-play communication increases the frequency of the efficient outcome. Alignment of incentives makes coordination through communication a reasonable explanation in games with multiple equilibria. For social dilemmas, scholars have also proposed behavioral explanations in line with reciprocity, fairness and a cost of letting others down (which can be fixed or depending on beliefs). Ellingsen and Johanssen (2004) for instance, develops a model that aim at explaining why some individuals reject positive amounts, why communication enhances outcomes and why equal splits are so pervasive in a hold-up game[2] by assuming preferences for equality and a fixed cost for being caught lying. Miettinen and Suetens (2008) measures guilt (through self-reported emotional reaction) when individuals do not honor their word in social dilemma games with pre-play communication. The measure of guilt is positively correlated with cooperation in a prisoner's dilemma game. Along the same lines, Charness and Dufwenberg (2006) provides a mechanism borrowed from psychological game theory through which guilt aversion leads to individuals to reciprocate if other trusted in a trust game: individuals may face a cost if they believe are letting others down. Gneezy (2010) argues that individuals may also posses a fixed cost of lying. In his experimental design Gneezy asks individuals to send a truthful and deceitful messages to another person about two possible splits of a pie. Then he makes subjects to play a dictator game with the same options in the message. He finds that a considerable proportion of his sample told the truth, but chose the selfish allocation in the dictator game, suggesting individuals may feel a cost of lying.

In a nutshell, our contribution is to integrate all these findings towards a better understanding of initiative taking. We start from the assumption that traits –reciprocal altruism and lying aversion- and incentives (context) determine leadership emergence. Precisely, by measuring reciprocal altruism

[2] The game consists of one party, the seller, deciding to invest some amount. If invested it gets multiplied, the the other party decides on a split of the multiplied amount. The seller has to decide whether to accept the split or reject it. If rejected, both inidviduals get nothing.

and lying aversion, we aim at assessing their effect on the decision to exhort others to cooperate in two different, but important contexts: a stag hunt game and a prisoner's dilemma game. We manipulate the incentives to take advantage of others, just to see to which extent reciprocity and lying aversion (and other personal traits) manifest on leaders under these different circumstances.

3 Experimental Procedure

The experiment was constructed to study the key drivers of leadership and measure the effectiveness of leadership activity. The underlying idea is that leadership is context specific. When a would-be leader merely needs to achieve coordination on an outcome consistent with selfish behavior, there is little personal risk to the leader and leadership is ubiquitous. In our framework, this is the case were rewards to loyalists are low. In contrast, when the situation is one where the leader must persuade others to override self-interest in choosing an action, only highly motivated individuals will assume the leadership role. In our particular context, motivation comes in the form of a desire for reciprocity and indifference towards lying. Neither of these traits has received much attention in the extant literature on leadership.

The design is structured to ascertain key personal characteristics such as reciprocity, lying aversion, altruism and risk aversion. It also measures standard leadership traits such as extroversion, agreeableness, internal locus of control and intelligence. Finally, we measure demographics such as gender, race and ethnicity. Age and experience might also correspond to leadership; however, our subject population of mainly undergraduates lacks much variation along these dimensions. The heart of the design consists of varying the context in which potential leadership emerges. We now turn to the details.

The experiment consisted of two treatments of two sessions each (96 subjects in total, 48 in each treatment). The instructions for the experiment were passed out to the participants and read aloud before the session began. A copy of the instructions is in the Appendix A. Participants did not interact with the experimenter, except to ask questions immediately after the instructions were read and before the experimental tasks began. The experimental currency was the Berkeley Buck ($) and the exchange rate was $12 per US$1. There are three parts to each experimental session. We describe them in the same order they were presented to participants in each session.

3.1 Part 1: Social Preferences and Lie Aversion

Part 1 was intended to elicit unconditional social preferences and a proxy for lying-aversion. The procedure in Part 1 was the same for all the treatments. It consisted of three blocks. In the first block, we elicited unconditional social preferences by asking the subjects to divide 10 tokens. The exact text presented in the computer screen was:

> Divide 10 tokens. Allocate a number of tokens to yourself (hold) and a number of tokens to the other participant (pass).
>
> A token is worth $X to you and $Y to the other participant. Please choose a division (total 10 tokens).
>
> Hold (1 token = $X):____
>
> Pass (1 token = $Y):____

Four different situations (values of X and Y) were presented to the subjects in the following order: (X=$1, Y=$1.25), (X=$1, Y=$1), (X=$1, Y=$0.67) and (X=$1, Y=$2). By varying the value of keeping versus passing each token, we can characterize the subject choices as either selfish or non-selfish. A subject was categorized as selfish if she kept all the tokens regardless of relative "prices;" otherwise a subject was labeled as non-selfish. This revealed preference elicitation procedure was first devised by Andreoni and Miller (2002) and subsequently extended by Fisman, Kariv and Markovits (2007).

Payouts for this part were calculated by randomly selecting one of the four allocation decisions. Subjects were randomly matched in anonymous pairs to execute the payouts dictated by that selected allocation. As a result, each participant received his/her value held and the value passed by his/her matched partner from the corresponding selected allocation. The matching was performed at the end of the experiment.

In the second block of Part 1 we elicited lying-aversion using a procedure similar to Gneezy (2010). Subjects faced two options featuring different divisions of $20. The options were: Option 1) keep $15 to him/herself and give $5 to other participant or Option 2) keep $5 to him/herself and give $15 to other participant. Subjects were asked to send a pre-codified message to another subject, who did not know which option corresponded to which set of payoffs. Participants could send a deceitful message reading "Option 1) will earn you [the subject the message was intended for] more money than Option 2);" or they could send a truthful message reading "Option 2) will earn you [the subject the message was intended for] more money than Option 1)." We randomized the order of Option 1) and Option 2) and used colors (Blue and Red) instead of numbers (Option Blue instead of Option 1), etc.) to avoid decisions based on mechanical ordering. Participants were anonymously matched in pairs at the end of the experiment. Subjects were told the message would be delivered to another randomly matched participant at that time, and the amount of money they both would get depended on this other subject's decision. Each subject received the payout from his/her own decision after observing the matched participant's message and the payout from the matched participant's decision after reading his/her own message.

We also asked subjects the probability their "advice" would be followed by the person with whom they would be matched. This permits us to distinguish between circumstances where the deceitful message was sent with an intent to deceive versus when it was sent to counteract the partner's skepticism about the veracity of the message. For instance, a subject wishing to offer helpful advice, but suspecting her partner will do the opposite of whatever message was sent, could only achieve her objective by sending the deceitful message rather than the truthful one. Thus, it seems important to distinguish white lies, deceitful messages intended to lead the partner to choose the higher payoff option, from black lies, deceitful messages sent with the intent to trick the partner into choosing the lower payoff option. In coding for lying aversion, we treat white lies as equivalent to truthful reports.

The procedure above potentially confounds lying aversion with altruism. A sufficiently altruistic subject who is not lying averse may still send a truthful message purely out of desire to be generous towards her partner. To untangle the two effects, we again follow Gneezy (2010) and implement a non-strategic version of the message game above. Here, every subject gives "advice" to a computer, which follows it with the same probability indicated by the subject in the game before. Subjects did not know the probability was going to be the same in both versions, they were only told in the instructions (see the instructions in the Appendix) the computer may execute their decision with "some" probability. Since lying to a computer does not carry the same moral stigma than lying to another person, choices in this round should purely reflect other-regarding preferences, to the extent they are present. We code a subject as a lying averse if they sent a truthful message or a white lie to a human and a selfish "message" to the computer. A subject is not lying averse otherwise, except for the cases in which truthful messages or white lies were accompanied by altruistic choices (12 of 96 (13%) subjects in total, 5 in the SH treatment and 7 in the PD treatment); because for these subjects, the differential payoffs between Option 1) and option 2) are not enough to elicit their aversion to lying. We exclude this subjects from the sample.

Payoffs in this latter procedure were determined exactly as in the previous one, except for the fact that the computer made all the decisions (reversal and matching). Figure 3 in Appendix B contains a screen shot of the interface for this portion of the experiment.

3.2 Part 2: Belief elicitation and treatment administration

In part 2 of the experiment, subjects played a one shot prisoner's dilemma followed by a stag hunt (SH treatment) or a prisoner's dilemma (PD treatment), depending on the treatment. This process was repeated 12 times, each with a different partner. The treatment was fixed over each session.

In the one-shot prisoners' dilemma (OSPD for future reference), subjects chose between Cooperate or Defect.[3] In the same screen in which subjects picked their action, they were asked to forecast how many other subjects would choose to cooperate. If a subject exactly predicted how many of the other 23 subjects in their session cooperated they would get \$8. Subjects lost \$1 times the (absolute) difference between their guess and the true figure.[4]

Half of the rounds of the OSPD were randomly selected for payment. In a selected round, each subject was randomly matched with a partner to compute payoffs. Similarly, forecasts for half of the rounds were compensated. These determinations were made at the end of the experiment. Thus, after each round, subjects received no feedback about their payoffs nor their forecasts.

This portion of the experiment provides a measure of reciprocal altruism. Subjects who cooperate when they forecast high cooperation from others are coded as reciprocal altruists.

In contrast with the other parts of the experiment, subjects were matched with another subject to participate in an interaction that would have immediate, rather than deferred, payoff consequences

[3] We labeled options as A or B, and randomized the link to Defect or to Cooperate. That is, in some rounds, A was the "Defect" action and in others it was the "Cooperate" action. The same was true for B. We also randomize the entries in the payoff matrix. We used different payoffs in each of the twelve rounds of the OSPD. We kept constant the net benefit of defection (equal to \$5) if the other cooperated and the net benefit of defection if the other defected (equal to \$4) to make it comparable to the PD game in Table (1). We also varied the order in which options (Cooperate or Defect) were presented. In some of the 12 rounds, the option equivalent to Defection was presented as the first row, and in others the option equivalent to Cooperation was presented in the first row, to avoid mechanical behavior. This ordering was random. These changes across rounds were intended to encourage participants to pay attention to the payoffs and the options in each round to avoid automatic responses.

[4] With this procedure, we elicit an statistic of the distribution (the median) of the number of participants who would cooperate. In theory, a risk neutral agent i chooses y to maximize

$$8 - \sum_{n=1}^{23} |y - n| p_i(n)$$

where $p_i(n)$ is the probability (belief) agent i assigns to n participants cooperating. The optimal choice $y^* = median_i(n)$. Therefore, y^* gives us information about the individual $i's$ beliefs, $p_i(n)$, about overall cooperation.

1\2	Defect	Cooperate
Defect	4, 4	**8**, 0
Cooperate	0, **8**	9, 9

1. SH

1\2	Defect	Cooperate
Defect	4, 4	**14**, 0
Cooperate	0, **14**	9, 9

2. PD

Table 1: Games in the two Treatment conditions

and feedback. In each round a subject would be matched with one other subject using a rotation matching protocol (Cooper, DeJong, Forsythe, and Ross 1996, Dal Bo 2005). This procedure divided participants in each session into two groups and then matched each subject in one group with one subject in the other group, without repetition. This ensured that any pair of subjects were matched at most once and that one subject was not matched with a participant in his/her own group. The goal of this procedure is to minimize strategic effects across rounds.

Each treatment consisted of two screens. In the first screen, participants observed a payoff matrix corresponding to the Coordination Game or Prisoner's Dilemma respectively, as in Table 1. On the left of each matrix, there was a chat box in which subjects could communicate for 30 seconds prior to making choices. Once the 30 seconds had elapsed, both subjects were directed to a second screen, again displaying the corresponding payoff matrix in Table 1, but now they had to choose whether to Defect or Cooperate simultaneously and without the opportunity to chat.[5]

The point of the chat portion of the design is to measure leadership, which we define to be the initiation of a plan to cooperate. Subsequent play allows us to assess the effectiveness (or lack thereof) of such initiatives. The two treatments vary the context in which leadership activities occur. In the SH treatment, cooperation is consistent with money maximizing play, though it does require a degree of trust between the two parties. In the PD cooperation is, of course, inconsistent with money maximizing play. Thus, a leader might persuade her partner to cooperate for social motives –mutual cooperation benefit both parties- or selfish motives –the leader stands to gain more from defection if her partner can be persuaded to cooperate.

The belief elicitation phase of the OSPD is designed to track how beliefs about overall cooperativeness might chance as a result of these interactions.

[5] As before we labeled A the Defect option and B the Cooperate option. Different from before, we did not randomize the order of these options as presented to subjects to make sure they are familiarized with the meaning of each option in case they want to communicate intentions to play to the other individual. Figure 4 in Appendix C shows the screens corresponding to this section.

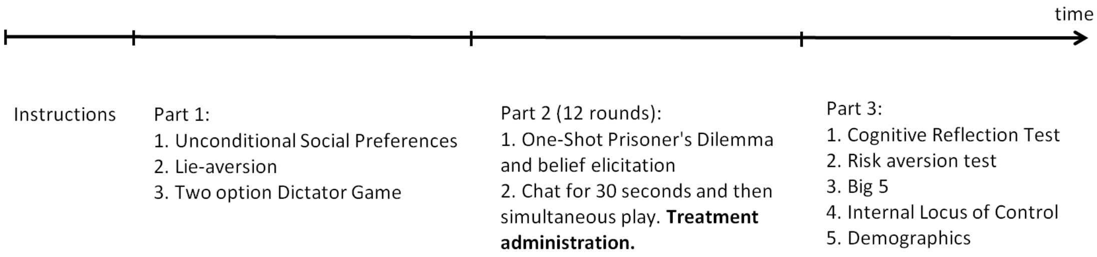

time
Instructions
Part 1:
1. Unconditional Social Preferences
2. Lie-aversion
3. Two option Dictator Game
Part 2 (12 rounds):
1. One-Shot Prisoner's Dilemma and belief elicitation
2. Chat for 30 seconds and then simultaneous play. **Treatment administration.**
Part 3:
1. Cognitive Reflection Test
2. Risk aversion test
3. Big 5
4. Internal Locus of Control
5. Demographics

Figure 1: Time line experimental procedure

3.3 Part 3: Questionnaires

Part 3 of the experiment consisted of a Cognitive Reflection Test (CRT, Frederick 2005), a Risk Aversion test (Holt and Laury 2002), a Big 5 personality test (John 1990), an Internal Locus of Control test (Rotter 1966) and a questionnaire about basic demographics (gender, major and ethnicity). From these tests, only that for Risk Aversion was incentivized. The goal of these questionnaires was to elicit traits that have been found relevant to leadership in the social psychology literature and to link them to initiative and cooperation in our setting.

Figure 1 shows the time line of the experiment.

We conducted 4 sessions during April and May of 2012 at the UC Berkeley Xlab. 96 UC Berkeley students from the Xlab subject pool participated in the experiment (95 undergraduates and 1 graduate). Sessions lasted approximately 1 hour and payoffs averaged US$16. Each participant took part in only one session. All treatments were programmed and conducted using z-Tree (Fischbacher 2007). Throughout the experiment we sought to ensure anonymity. Participants were separated in workstations and no communication was allowed other than that feasible through the chat-box for the 30 seconds in each round.

4 Hypotheses

Consider the monetary payoffs in Table 1, corresponding to the case where the state does and does not reward loyalists (stag hunt and prisoner's dilemma games). Reciprocal altruism is the private intrinsic reward of joining the demonstration if and only if the other demonstrate as well. In Table 2, the parameter $\rho_i \geq 0$, $i = 1, 2$ represents this extra gain. For the low reward game, a higher ρ_i

1\2	Defect	Cooperate
Defect	$4,4$	$\mathbf{8},0$
Cooperate	$0,\mathbf{8}$	$9+\rho_1, 9+\rho_2$

1. SH

1\2	Defect	Cooperate
Defect	$4,4$	$\mathbf{14},0$
Cooperate	$0,\mathbf{14}$	$9+\rho_1, 9+\rho_1$

2. PD

Table 2: Games in the two Treatment conditions

reinforces the incentives for mutual cooperation. In the prisoner's dilemma however, a higher ρ_i may flip incentives to join the demonstrations –cooperation may be a best response to cooperation.

Everyone can only benefit from others demonstrating. In the "low rewards to loyalists" regime (SH treatment), everyone joins the manifestations. In the "high rewards to loyalists" regime (PD treatment), however, not everyone demonstrates out on the streets if others do so. Some individuals may stay at home (if ρ_i is very low) and "free-ride" on others. If communication is costless, everyone, including these "free-riders," may try to exhort others to demonstrate. No exhortation to cooperate may be credible in this regime, unless individuals suffer a cost of lying. If that is the case, the ones who are not willing to honor their word remain silent. Our first hypothesis is that, on average, anyone takes the initiative to convince others to go out to the streets in the "low rewards to loyalists" regime. In the "high rewards to loyalist"regime only special individuals do so: Those with high reciprocal altruism exhort others to cooperate more often than individuals with low reciprocal altruism and individuals with high lying aversion do it less often than those with low lying aversion. Thus, there is an interaction effect of reciprocal-altruism and lying aversion. Individuals with high reciprocal altruism are willing to demonstrate if they are optimistic enough that others will do so. Lying aversion for those individuals should not affect their decision to lead. If their reciprocal altruism is instead low, then high lying averse individuals initiate less often than those for whom lying is not too costly.

When loyalists are rewarded for not demonstrating, motivated individuals (those with high reciprocal altruism, ρ_i) emerge more often. Those who emerge demonstrate more often. Our second hypothesis is that, the regime falls more often when there is an individual emerging than when no one emerges. This effect of emergence on successful demonstrations is less pronounced when the regime that does not compensate loyalists.

Overall, we should observe that a great majority should end up adhering to demonstrations when rewards for loyalty are low. When they are high however, only special individuals, with high reciprocal altruism and high lying aversion do so. This is our third hypothesis.

$i\backslash j$	Defect	Cooperate
Defect	d_t, d_t	a_t, b_t
Cooperate	b_t, a_t	$c_t + \rho_i, c_t + \rho_j$

Table 3: OSPD game

5 Experimental results

In this section we first describe the constructs created from the decisions made in the experiment linking them to our theoretical specification. Then we show the results.

5.1 Constructs

5.1.1 Initiative

Our proxy for initiative comes from the first message sent suggesting mutual cooperation. In our data, this first message can be classified as: Bold Suggestion of Cooperation ("Bold Suggestion"), Proposal of Cooperation ("Proposal"), "No Suggestion" or "Suggestion of Defection." We categorize any initiative as "Bold Suggestion" if it involves compromise such as (the label "A" represents defection and "B" cooperation): "I'll choose B" or "We both should choose B". Initiative in the form of "Proposal of Cooperation" is taken to point out or to reference the cooperative outcome; for example: "B and B?" or "Shall we both go B?". No Suggestion is simply not sending a message about cooperation during the chat stage and "Suggestion of Defection" comprises the ones who first suggested to defect. "Suggestion of Defection" is infrequent (4 of 576 (0.7%) in the SH treatment and 29 of 576 (5%) in the PD treatment). More examples of each category are in the Appendix. We code as initiative any "Bold Suggestion" or "Proposal" and as no initiative otherwise.[6]

5.1.2 Reciprocal altruism: ρ

We estimate reciprocal altruism from the decisions in the one-shot prisoner's dilemma game (OSPD) preceding the treatment conditions. We first estimate a proxy for each individual's reciprocal altruism and then we classify them as either low or high depending on whether they are above or below the median.

Let us consider the prisoner's dilemma game (OSPD) in round $t = 1, ..., 12$ with utilities as in Table 3.

[6] We also tried excluding the "Suggestion of Defection" instances from the analysis. All the result hold as well.

Let us denote σ_{jt} the belief about individual j cooperating in round t from i's perspective. In the OSPD game, individual i cooperates in round t if and only if

$$\begin{aligned} E[U_{it}(C;\rho_i)] &> E[U_{it}(D;\rho_i)] \qquad (1) \\ \iff (1-\sigma_{jt})b_t + \sigma_{jt}(c_t+\rho_i) &> (1-\sigma_{jt})d_t + \sigma_{jt}a_t \\ \iff \rho_i &> \frac{(1-\sigma_{jt})}{\sigma_{jt}}(d_t-b_t) + (a_t-c_t) \equiv \rho_t^*(\sigma_{jt}) \end{aligned}$$

or individual i defects in round t if and only if $\rho_i \leq \frac{(1-\sigma_{jt})}{\sigma_{jt}}(d_t-b_t) + (a_t-c_t) \equiv \rho_t^*(\sigma_{jt})$. From the data we observe the decision (Cooperate or Defect) in the OSPD and we elicit σ_{jt} (a_t, b_t, c_t and d_t are known). With σ_{jt} we are able to compute $\rho_t^*(\sigma_{jt})$ for each round t. Let us denote $c_{it} = C$ if individual i cooperates in round t and $c_{it} = D$ if defects, then we may deduce that, for $t = 1, ...12$,

$$\rho_i \in [\max_t\{\rho_t^* | c_{it} = C\}, \min_t\{\rho_t^* | c_{it} = D\}]. \qquad (2)$$

Expression (2) indicates that a rational individual whose preferences are described in (1) must have a reciprocal altruism parameter at least as high as the one who leaves him indifferent about cooperating under the most pessimistic beliefs, and at most as high as the one which makes him defect under the most optimistic beliefs about other's cooperating.

This is assuming that all the decisions are rational, according to this model. We first need to check to which extent individuals behave under these primitives. To that end, we compare all the decisions made in pairs of rounds (12 rounds, 66 pairs in total) to see which pairs are consistent. By consistent we mean that if an individual chooses to Cooperate (Defect) given beliefs σ_{jt} in round t then she must choose to Cooperate (Defect) in round $t' \neq t$ if the beliefs $\sigma_{jt'}$ are more optimistic (pessimistic), $\sigma_{jt'} \geq \sigma_{jt}$ ($\sigma_{jt'} \leq \sigma_{jt}$).

In the SH treatment 23 of 48 (48%) individuals are rational under this model (that is all the 66 pairs of decisions are consistent) and 38 of 48 (80%) show fewer than 5 of 66 pairs of inconsistent decisions. In the PD treatment the account is similar: 18 of 48 (38%) individuals are fully rational and 33 of 48 (69%) show fewer than 5 of 66 pairs of inconsistent decisions.

For individuals who show no inconsistent decisions, the interval in (2) can be created directly. For the remaining subjects, we exclude the decisions (from the 12) that are most inconsistent with others. For instance, participant S1-11 (session 1 subject number 11) in the PD treatment presents 3 of 66 pairs with inconsistent decisions. One decision (made in round 10) is inconsistent with three others (those made in rounds 7,8 and 11). Once we exclude that decision, all the eleven remaining choices are consistent with each other, so interval (2) can be created. We follow this procedure for every participant who have at least one inconsistent pair. Overall and after this procedure of excluding the inconsistent decisions, 42 of 48 participants (88%) in the SH treatment have 10 or more consistent decisions, and 1 of 48 (2%) has the minimum of 7 consistent decisions. In the PD treatment 38 of 48 (79%) participants have 10 or more consistent decisions, and 4 of 48 (8%) have the minimum of 8 consistent decisions. We use only these consistent decisions for each subject to compute the interval (2).

There are some participants who either always defect or always cooperate in the OSPD game when considering only consistent decisions. For these, we could compute only one of the boundaries of interval (2).[7] As might be expected, defection is more common than cooperation in this game. From all the subjects, 31 of 96 (32%) always defect (no observations to estimate the lower bound) and 15 of 96 (16%) always cooperate (no observations for the upper bound). We could have pursued estimating ρ_i from strictly within interval (2), which would have required assumptions on the missing boundary for all these 46 individuals. We follow a different approach. In order to exploit the information in the data and minimize the number of ad hoc assumptions, we estimate the reciprocal-altruism parameter ρ_i as the upper bound. This estimation procedure has one important implicit assumption: Individuals who defect in the OSPD when they are more optimistic are more likely to be low reciprocal altruists than the ones who defect when they are relatively more pessimistic. In the appendix we explore other constructs that yield to similar results.

We discretize our reciprocal altruism construct (high or low) to avoid problems derived from outliers (subjects whose decisions in the OSPD are either always defect or always cooperate and are rational given their beliefs). The discretization is as follows. For each treatment, we classify as a low (high) reciprocal altruist if the upper bound of interval (2), computed as described in this section, is below (above) the median value of this upper bound of all the subjects in the sample for that treatment. Individuals who cooperated in every round are coded as high reciprocal altruists. This procedure yields to 25 of 48 (52%) individuals classified as low reciprocal altruists and 23 of 48 (48%) individuals classified as high reciprocal altruists in the SH treatment. In the PD treatment the equivalent figures

[7] The reason is that when an individual always defected the set $\{\rho_t^*|c_{it} = C\}$ is empty, so the $\max_t\{\rho_t^*|c_{it} = C\}$ does not exist. Similarly, for an individual who cooperated in all the 12 rounds, the set $\{\rho_t^*|c_{it} = D\}$ is empty, so the $\min_t\{\rho_t^*|c_{it} = D\}$ does not exist.

are 20 of 48 (42%) and 28 of 48 (58%).

5.1.3 Lying-aversion: λ

Our construct for lying aversion comes from the second block of Part 1, in which participants can send a truthful or a deceitful message to a randomly matched partner. As in Gneezy (2010) we classify a given individual as lying-averse if he/she sends a truthful message and declares the other would follow with at least 50% chance, but choose the option that gave him/her more money in the two-option dictator game that follows. We extend this definition to include subjects who believe the partner will follow with less than 50% chance. Sending the deceitful message is intended to help the partner ("white lie"). We code subjects as "lying-averse" if they sent a truthful message or a "white lie," but chose the selfish expected payoff in the subsequent equivalent game with the computer. Otherwise they are coded as not lying averse. In the SH treatment 19 of 48 (40%) participants are coded as lying-averse and 29 of 48 (60%) are coded as not lying averse. In the PD treatment 15 of 48 (31%) are coded as lying-averse and 33 of 48 (69%) are coded as not lying averse under this criterion. This mechanism provides us with a lying aversion construct comparable to Gneezy's (2010). [8]

5.1.4 Unconditional social preferences, personality traits and demographics

We also elicit unconditional social preferences and perform a battery of tests to measure personality traits and demographics. By means of the 4 menus in block one of Part 1, we are able to distinguish 27 of 96 (28%) perfectly Selfish subjects (they keep everything in each menu).[9]

The first test is the Cognitive Reflection Test developed by Frederick (2002). It consists of three questions and aims at measuring cognitive ability. Individuals were given 5 minutes to answer the test, after which the screen disappears. The score is one point for each correct answer, so 0 is the minimum and 3 is the maximum. Immediately after we administered a short version of the risk-aversion test in Holt and Laury (2002). It consisted of 4 alternatives, presented in order, in which each individual had to choose among two lotteries, one riskier than the other. Table 5 Appendix D shows an screen shot. A risk averse individual should start by choosing the safe one and may switch to the riskier one when it becomes attractive enough. 4 participants presented inconsistent choices (switched more than

[8] Notice there are subjects who send the truthful message or a "white lie" who choose the altruistic outcome afterwards. We code these subjects as not lying averse. In the appendix we show the result holds if we exclude this subjects from the sample.

[9] Andreoni and Miller (2002), found 23% of their subjects can be classified as perfectly selfish and Fisman et al (2007) found that was the case for 26.3% of their sample.

once), 3 of the 92 (4%) remaining were classified as risk loving (they never chose the safe lottery) and 16 of 92 (17%) as extremely risk averse (they never switched to the risky lottery).

We conducted a Big 5 personality test (John 1990). It consists of 44 questions to characterize individuals based on 5 personality traits: Extroversion (or Extroversion), Agreeableness, Conscientiousness, Neuroticism and Openness. Each question asks about own perception of personality attributes. These attributes have been found to be strongly correlated with leadership (Judge et al. 2002). Extroversion has been associated to leadership emergence (Bass 1990, Gough 1990) mainly because leaders emerging from leaderless groups are more active, energetic, not silent and assertive (Gough 1988). Agreeable individuals tend to be cooperative and sensitivity, tact and altruism seem to be the defining features of an agreeable personality. The evidence on the direction of the relationship between leadership and agreeableness is however not clear. Cooperativeness tends to be positively related to leadership, but sensitivity and tact are more likely to be related to modest individuals, who do not usually emerge as leaders (Bass 1990). Conscientiousness is related to task oriented behavior which is in turn associated with effective leadership (Barrick and Mount, 1991),). Two factors, high self-esteem and high self-confidence, are traits find in most of the trait based studies on leadership. These two characteristics are related to low Neuroticism (Bass 1990). Openess main components are creative and divergent thinking, both of which have been positively related with effective leadership (Sosik, Kahai and Avolio, 1998).

Our last test is version of the Internal Locus of Control test developed developed in Rotter (1966). It consists of 13 questions, with two statements each, one indicating that people have no control over a certain hypothetical event, and the other suggesting the opposite. Individuals with a higher score on this test (more responses associated with the “control over events” alternatives) have been found to be high-achievers (ambitious and task oriented) and directed by own beliefs, rather than extraneous advice (Rotter 1970).

Finally, we ask for demographic characteristics such as gender and ethnicity.

5.1.5 Constructs for initiative and cooperation

After an individual takes the initiative, the matched partner can either reply by agreeing on the suggestion to cooperate (join the demonstrations), or say nothing (or say something unrelated to cooperation or defection). We code as 1 if the former happens and zero otherwise. Cooperation is equal to 1 if the participant chooses to join and zero otherwise, in each treatment.

	CG N	CG mean	PD N	PD mean	diff	se	p
Reciprocal Altruism (ρ)	48	0.48	48	0.58	-0.10	0.10	0.31
Lying Aversion (λ)	48	0.40	48	0.31	0.08	0.10	0.40
Selfish	48	0.63	48	0.63	0.00	0.10	1.00
InternalLocusofControl	48	6.67	48	6.25	0.42	0.48	0.39
Extraversion	48	3.11	48	3.22	-0.11	0.17	0.50
Agreeableness	48	3.65	48	3.69	-0.05	0.11	0.66
Conscientiousness	48	3.43	48	3.49	-0.06	0.14	0.67
Neurotism	48	2.82	48	2.79	0.03	0.14	0.83
Openness	48	3.48	48	3.49	-0.01	0.12	0.93
ScoreCRT	48	1.48	48	1.35	0.13	0.24	0.60
RiskAversion	45	3.38	47	3.49	-0.11	0.21	0.59
female	48	0.73	46	0.67	0.06	0.10	0.56
Asian	48	0.71	48	0.65	0.06	0.10	0.52
White	48	0.19	48	0.23	-0.04	0.08	0.62
OtherEthnicity	48	0.10	48	0.13	-0.02	0.07	0.75
Initiate	576	0.58	576	0.41	0.17	0.03	0.00
Cooperate	576	0.82	576	0.39	0.42	0.03	0.00

Table 4: Summary statistics

Table 4 provides a summary of the variables just described. In what follows we use the term "cooperation" to represent individuals adhering to the demonstrations.

5.2 Initiative and types

Overall, initiative is pervasive in both games, but initiative is significantly (chi-squared p-value=0.000) more common in the SH treatment, in which the rewards for loyalists are low. 332 of 576 (58%) individual choices involve suggestion of cooperation in the SH treatment, while only 234 of 576 (41%) do so in the PD treatment. In our framework, rewards to loyalists indeed matter for the frequency a leader emerges to take people out to the streets.When honoring own's word and following is optimal, more individuals take the initiative than when individuals may succumb to the temptation of staying home to enjoy the high rewards the incumbent regime concedes to loyalists.

Our main result links initiative (exhorting others to demonstrate) to reciprocal altruism and lying-aversion. Reciprocal altruism and lying-aversion matter for initiative in the PD, but not in the SH treatment. Precisely, in the PD treatment, higher lying averse types initiate less often if they are also low reciprocal altruists, but more often if they are high reciprocal altruists.

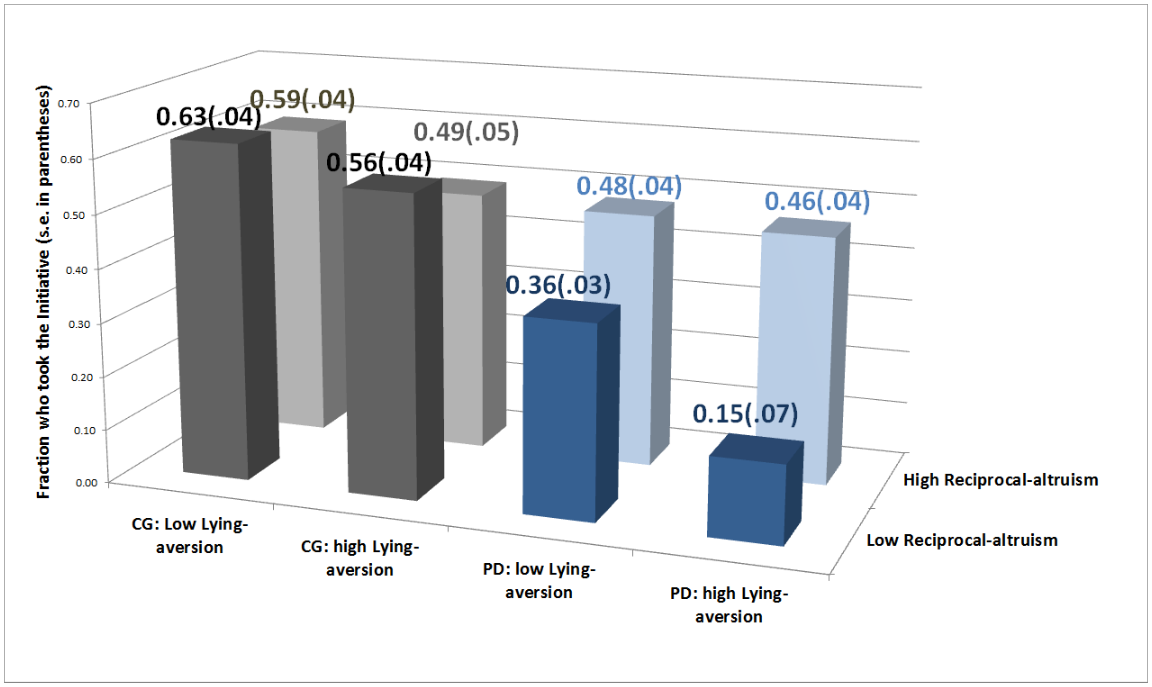


Figure 2: Rates of initiative by Treatment for each type (ρ, λ)

Support for this result comes from Figure 2 and Table 5. Figure 2 shows the rate of initiative separated by each pair of types. In the SH treatment, no matter the type of reciprocal altruism and lying aversion, initiative is around 60% for each of them. In the PD treatment however, only 15% of the individuals coded as low reciprocal altruist and high lying aversion take the initiative. In theory, these individuals are less prone to initiate, because the prospect of not adhering to her word when initiating looms larger than when agreeing. High reciprocal altruists types take the initiative 50% of the time in the PD treatment. Individuals who in theory care little about cooperation and lying doesn't seem to be a problem, low reciprocal altruism and low lying aversion types, take the initiative 33% of the time. Notice also from the picture that, overall, high reciprocal altruism is related to more initiative, and high lying aversion is related to less initiative.

Table 5 shows the same pattern through a reduced form regression of initiative on our measures of reciprocal altruism and lying aversion and the interaction between them. Including the interaction is important because it captures the nuanced effect of lying aversion on initiative. The first two columns in Table 5 exhibit the results for the SH treatment. The first column shows the estimations for the parsimonious specification that includes only our constructs for reciprocal altruism and lying aversion, and the second includes all the controls. Adding the controls does not make any of our estimations for the coefficients of reciprocal altruism nor lying aversion significant, but it shows that Internal Locus of Control, Agreeableness, risk aversion and ethnicity are significantly related to initiative in the SH treatment. The last two columns show the results for the PD treatment. The coefficient for lying aversion is significant in the third specification and the three coefficients (the one for lying aversion, the one for reciprocal altruism and the one for the interaction between them) are significant at conventional levels for fourth specification. The coefficient on lying aversion is negative, implying that lying averse individuals initiate less, when they are also low reciprocal altruists. In this case, we expect that individuals evaluation of the costs and benefits of leading are determined by individual traits. In particular, for individuals whose intrinsic motivation to mutually cooperate is low (low reciprocal altruists) initiate less the higher is their lying aversion. High reciprocal altruism is positively related to initiative with and without controls, and motivated individuals, those who have high reciprocal altruism and high lying aversion, take the initiative more often. Moreover, the fourth column suggests that Conscientiousness and Openness are positively related to initiative and Agreeableness, score in the CRT and being female are negatively related.

These results offer evidence to support our first hypothesis: individual traits in the form of reciprocal altruism and lying aversion matter little in the SH treatment, but they do matter in the PD treatment. High lying aversion individuals initiate less, high reciprocal altruistic individuals initiate more, especially if they are also high lying averse. This is consistent with our first hypothesis on the direction reciprocal altruism and lying aversion relate to initiative.

5.3 Initiative, agreement and cooperation

As expected, individual participation in revolts is more frequent in the SH ("low rewards to loyalist" regime) than in the PD treatment ("high rewards to loyalists" regime). In 470 of 576 (82%) cases individuals cooperated in the SH treatment. In the PD treatment, this was the case in 227 of 576 (39%) decision instances. Looking at group outcomes, in the SH treatment in 202 of 288 (70%) game-rounds the group ended up with both individuals cooperating (overthrowing the regime) while in 20 of 288 (7%) it ended up with both defecting. In the PD treatment, the equivalent figures are 67 of 288 (23%) and 128 of 288 (44%).

	(1)	(2)	(3)	(4)
	CG	CG	PD	PD
	Pr$\{m_i$=1$\}$	Pr$\{m_i$=1$\}$	Pr$\{m_i$=1$\}$	Pr$\{m_i$=1$\}$
Low Reciprocal-A. High Lying-A.	-0.17	-0.05	-0.69***	-1.09***
	(0.25)	(0.26)	(0.16)	(0.31)
High Reciprocal-A. Low Lying-A.	-0.09	-0.00	0.30	0.47**
	(0.24)	(0.25)	(0.21)	(0.22)
High Reciprocal-A. High Lying-A.	-0.34	-0.26	0.27	0.73***
	(0.29)	(0.33)	(0.24)	(0.27)
Selfish		-0.26		0.22
		(0.24)		(0.17)
InternalLocusofControl		-0.09*		-0.02
		(0.05)		(0.05)
Extraversion		0.12		0.04
		(0.11)		(0.14)
Agreeableness		0.39*		-0.34**
		(0.22)		(0.17)
Conscientiousness		0.11		0.31**
		(0.18)		(0.14)
Neuroticism		0.14		0.13
		(0.18)		(0.14)
Openness		-0.29		0.62***
		(0.20)		(0.13)
ScoreCRT		-0.03		-0.16*
		(0.10)		(0.08)
RiskAversion		-0.35***		-0.04
		(0.12)		(0.10)
female		0.01		-0.47*
		(0.24)		(0.26)
Asian		0.55***		0.03
		(0.20)		(0.29)
White		0.18		-0.51
		(0.39)		(0.43)
_cons	0.32*	0.19	-0.36**	-2.13**
	(0.17)	(1.52)	(0.14)	(0.89)
N	576	540	576	540
pseudo R^2	0.006	0.074	0.029	0.078

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Reduced form of initiative on reciprocal altruism, lying-aversion and controls

5.3.1 Initiative and cooperation

Initiative by itself does not guarantee effective leadership, because individuals may fail to follow suit. In that case one might worry that the previous results do not shed light on a useful kind of leadership. For leadership to be effective in overthrowing the regime then, individuals willing to mutually cooperate ought to believe that others' exhortation to cooperate must carry some truth; some leaders have to be willing to honor their word. Consistent with this we find that, on average, cooperation is more frequent among individuals who take the initiative than among those who do not, in both contexts. That is, in the SH treatment, among all the decisions in which an individual does not take the initiative, in 70% (170 of 244) of them he/she cooperates; whereas among all the individuals who do take the initiative, 90% (300 of 332) end up cooperating (chi-squared p-value=0.000). The pattern is similar in the PD treatment: 29% (98 of 342) of those who do not take the initiative cooperate, while 55% (129 of 234) does among those who take the initiative (chi-squared p-value=0.000). Table 8 in the Appendix E shows this result when clustering standard errors at the individual level and when we control for the other elicited covariates.

If at least some initiators intend to cooperate, then we may expect that some of those who do not take it may be willing to cooperate, especially if their reciprocal altruism is high. We observe that initiative induces others to cooperate only in the PD treatment. In the PD treatment, among all of the instances in which a given individual does not suggest cooperation, 35% (121 of 342) of the partners end up cooperating, while in instances in which a subject does suggest cooperation, 45% (106 of 234) of the partners actually cooperates (chi-squared p-value=0.017). In the SH treatment however, the proportion of partners cooperating after observing initiative by other does not change significantly. With no initiative, 82% (201 of 244) of the partners cooperate, while with initiative, 81% (269 of 332) of the partners cooperate (chi-squared p-value=0.679). Table 9 in the Appendix F shows this result when clustering standard errors at the individual level and when we control for the covariates. This result highlights an important feature that differentiates these two contexts: In the PD treatment, initiative is a costly signal of intention to cooperate that may convince others to follow suit, while in the SH treatment, initiative is taken by almost everybody (it is not costly) because there is no strategic conflict in the decision to cooperate.

Here the evidence for our second hypothesis is mixed. Initiative matters for initiators' cooperation in both games. However, exhortation (to induce cooperation in others by taking the initiative) is only statistically significant in the PD treatment, as our second hypothesis suggests.

5.3.2 Agreement and cooperation

What is the role of agreement? Initiative seems to matter for cooperation; is a pair that comes to an agreement more likely to end up in cooperation? Before we move to our third hypothesis on the relation between types, initiative and cooperation, we observe that agreement increases the chances of cooperation in both games. Moreover, cooperation is more frequent within individuals who take the initiative than within individuals who agree, in those games in which the pair comes to an agreement.

In the SH treatment, cooperation obtains in 43% (46 of 106) of the instances if agreement is not reached (because either no one took the initiative or initiative was taken, but the other party does not explicitly agree to adhere), but in a 90% (424 of 470) if agreement obtains (chi-square p-value=0.000). For the PD treatment, the pattern is similar, although perhaps more stark. Individuals in groups that do not observe agreement after the communication round cooperate in 13% (33 of 254) of the cases, while after agreement they do it in 60% (194 of 322) of the cases (chi-square p-value=0.000).

Individuals whose counterpart suggested to cooperate can decide to explicitly agree or not (by saying nothing or something else). Among the interactions (groups) in which one individual takes the initiative and the other agrees, 95% (165 of 174) of the initiators cooperate and 83% (144 of 174) of the ones who agree does, in the SH treatment. In the PD treatment, 65% (93 of 143) of the individuals taking the initiative cooperate after the counterpart's agreement, while 57% (82 of 143) of the ones who agree does (chi-squared p-value=0.000).

5.3.3 Types, agreement and cooperation

In the high rewards regime, some special types take the initiative more often. If these types also have a different speech, this may generate a selection bias on the coefficient of reciprocal altruism and lying aversion in a reduced form regression of cooperation on initiative and reciprocal altruism and lying aversion. In order to minimize this endogeneity problem, we provide evidence on cooperation by groups. Table 6 shows a normal probability model of both individuals cooperating on the characteristics of the members of that group. The first two columns exhibit the results for the SH treatment. The first column shows the correlations between cooperation unconditional on the pair coming to an agreement and the second, cooperation conditional on agreement. The third and fourth columns show the same specification for the PD treatment.

We find that high lying aversion is related to more cooperation in general and in pairs that come to an agreement only in the PD treatment. Support for this result can be seen from the coefficients in

	CG	CG	PD	PD
	Pr{Both C}	Pr{Both C$\|$Agreement}	Pr{Both C}	Pr{Both C$\|$Agreement}
At least 1 high Rec-Alt.	0.11	0.19	0.40	0.37
	(0.22)	(0.28)	(0.37)	(0.42)
At least 1 high L-A	-0.15	0.08	0.94***	1.23***
	(0.20)	(0.25)	(0.25)	(0.31)
GroupPerfectlySelfish	-0.24	0.37	-0.72***	-1.00***
	(0.18)	(0.26)	(0.20)	(0.27)
GroupILoC	-0.37***	-0.39***	0.07	0.05
	(0.08)	(0.11)	(0.08)	(0.10)
GroupExtraversion	-0.68***	-0.53**	-0.33	-1.00**
	(0.18)	(0.24)	(0.28)	(0.42)
GroupAgreeableness	1.04***	0.93**	1.94***	2.03***
	(0.33)	(0.45)	(0.36)	(0.44)
GroupConscientiousness	-0.57*	-0.71*	0.49	-0.21
	(0.31)	(0.42)	(0.30)	(0.41)
GroupNeuroticism	-0.28	-0.30	0.47	0.02
	(0.26)	(0.32)	(0.29)	(0.35)
GroupOpenness	0.64**	1.39***	0.13	0.66
	(0.28)	(0.41)	(0.34)	(0.51)
GroupScoreCRT	0.17	0.03	-0.21	-0.28
	(0.16)	(0.21)	(0.19)	(0.23)
GroupRiskAversion	-1.03***	-0.94***	0.04	0.13
	(0.24)	(0.31)	(0.21)	(0.28)
Groupfemale	0.08	0.15	0.41*	0.72**
	(0.18)	(0.24)	(0.22)	(0.29)
GroupAsian	-0.07	-0.14	-0.07	0.15
	(0.23)	(0.29)	(0.28)	(0.35)
GroupWhite	-0.65**	-0.45	0.12	0.08
	(0.27)	(0.35)	(0.35)	(0.42)
_cons	3.78	1.75	-11.78***	-8.13***
	(2.44)	(3.23)	(2.49)	(2.89)
N	273	222	286	161
pseudoR^2	0.238	0.256	0.313	0.336

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Group members' reciprocal altruism and cooperation, by treatment

Table 6. In the SH treatment, the effects of reciprocal altruism and lying aversion are not statistically different from zero, while in the PD treatment, at least one lying averse type is significantly related to more cooperation. These specifications also show some other variables related to cooperation. In the SH treatment, the average members' Internal Locus of Control, Extroversion, Conscientiousness, Risk Aversion and number of white individuals are negatively related to cooperation; The average of members' score on Agreeableness and Openness are positively related to cooperation. In the PD treatment, the members' score on the number of unconditionally selfish individuals and Extroversion have a negative relation with cooperation, and the members' score on Agreeableness and the number of females have a positive relation with cooperation.

This result is in line with our third hypothesis: types, reciprocal altruism and lying aversion, are related to cooperation only in the PD treatment.

5.3.4 Timing of initiative and agreement

There are other aspects of the conversation other than initiative and agreement that may also influence adherence. One of these aspects is how long (in seconds) did it take for the first individual to suggest cooperation, and the time (in seconds) it took to the non initiator to agree. We find that conditional on agreement, fast initiative and fast time of agreement increases the chances of cooperation in both treatments.

Conditional on agreement, fast initiative (defined as the time at which a given individual took the initiative being smaller than the median time overall groups) interacted with fast agreement (defined in the same way, but now with respect to the time it takes a non initiator to agree, if he/she actually does it) is positively related with more cooperation. The result holds even when we control for the number of reciprocal altruists and lying-averse individuals and the other measures we have used throughout. Table 7 shows the estimates for each treatment with and without controls.

6 Conclusion

In leaderless groups, coordination requires a special member to take the initiative to foster cooperation. We find that the traits that matter for leadership depend on the context. When the incentives to stay at home when others are demonstrating to overthrow a "malevolent" regime are low, no special talent is required from leaders, anyone initiates and anyone follows suit. When the regime rewards those loyalists who do not participate in demonstrations, leaders are indeed special. Individuals who exhort others to

	CG: Pr{Both C |Agreement}	CG: Pr{Both C |Agreement}	PD: Pr{Both C |Agreement}	PD: Pr{Both C |Agreement}
Slow Initiative, Fast Agreement	0.52	0.53	0.59*	0.30
	(0.35)	(0.46)	(0.35)	(0.43)
Fast Initiative, Slow Agreement	0.20	0.31	0.18	0.46
	(0.33)	(0.42)	(0.35)	(0.43)
Fast Initiative, Fast Agreement	0.59*	0.65	0.85**	0.98**
	(0.31)	(0.40)	(0.33)	(0.45)
_cons	0.43*	4.72	-0.74***	-8.19***
	(0.25)	(4.09)	(0.27)	(3.06)
CONTROLS	NO	YES	NO	YES
N	172	164	142	142
pseudo R^2	0.027	0.314	0.049	0.380

Standarderrorsinparentheses
$p < 0.10$,**$p < 0.05$,***$p < 0.01$

Table 7: Timing of initiative and agreement conditional on agreement, by treatment

demonstrate out posses higher intrinsic value for mutual cooperation and moreover a higher inclination to honor their word. Thus, this traits have complementary effects on initiative. If individuals have high motivation for mutual cooperation, being lying averse increases the chances to initiate a revolt. If individuals have low intrinsic motivation for mutual cooperation, then high lying aversion makes them even less likely to initiate. If the society consists of more honest mutually cooperative individuals, the chances of demonstration and that the regime actually falls are higher than when most of the population do no specially care about mutual cooperation, no matter how honest they are.

References

Andreoni, James and Miller, John H. "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence." Economic Journal, May 1993, 103(418), pp.570- 85.

Arbak, E., Villeval, M.-C., 2011. Endogenous leadership. Selection and influence. IZA Discussion Paper No. 2732.

Bass, B.M. (1990). Bass and Stogdill's Handbook of Leadership. New York: Free Press.

Cooper, Russell; Dejong, Douglas V.; Forsythe, Robert and Ross, Thomas W. "Cooperation without Reputation: Experimental Evidence from Prisoner's Dilemma Games." Games and Economic Behavior, February 1996, 12(2), pp. 187-218.

DeRue, D. S., Nahrgang, J. D., Wellman, N., & Humphrey, S. E. (2011). Trait and behavioral theories of leadership: An integration and meta-analytic test of their relative validity. Personnel Psychology.

Frederick, S. "Cognitive Reflection and Decision Making". Journal of Economic Perspectives. Volume 19, Number 4, Fall 2005, pp. 24-42.

Gneezy, U. "Deception: The role of consequences," American Economic Review, March 2005, 384-394.

Galton F, Eysenck HJ. (1869). Hereditary genius: London, England: Macmillan.

Güth W., & M. V. Levati, M. Sutter and E. van der Heijden (2006). "Leading by example with and without exclusion power in voluntary contribution experiments," Papers on Strategic Interaction 2006-35, Max Planck Institute of Economics, Strategic Interaction Group.

John, O. P., Naumann, L. P., & Soto, C. J. "Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement, and Conceptual Issues." In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), Handbook of personality: Theory and research (pp. 114-158). New York, NY: Guilford Press, 2008.

Judge TA, Bono JE, Ilies R, Gerhardt MW. (2002). Personality and leadership:Aqualitative and quantitative review. Journal of Applied Psychology, 87, 765–780.

Judge, T.A., Piccolo, R.F. and Ilies, R. (2004). "The Forgotten Ones? The Validity of Consideration and Initiating Structure in Leadership Research" Journal of Applied Psychology. 89(1) 36-51.

Krause, J., & Ruxton, G. D. (2002). Living in groups. Oxford, England: Oxford University Press.

Kreps, D. “Corporate Culture and Economic Theory,” in James Alt and Kenneth Shepsle, eds., Perspectives on Positive Political Economy. New York: Cambridge University Press, pp. 90-143, 1990.

Kreps, David M.; Milgrom, Paul; Roberts, Johnand Wilson, Robert. "Rational Cooperation in the Finitely Repeated Prisoner’s Dilemma." Journal of Economic Theory, August 1982,27(2), pp. 245-52.

Ledyard (1994) Public Goods a Survey of Experimental Research. Kagel and Roth eds. Handbook of experimental economics.

March J.G. and H. Simon (1958) Organizations. New York. Wiley.

Nosenzo, D., Sefton, M., 2011. Endogenous move structure and voluntary provision of public goods: theory and experiment. Journal of Public Economic Theory 13 (5), 721-54.

Rivas, M.F., Sutter, M., 2011. The benefits of voluntary leadership in experimental public good games. Economics Letters 112 (2), 176-8.

Stogdill, R.M. (1963), Manual for the Leader Behavior description questionnarie–for XII. Columbus: Ohio State University, Bureau of Business Research.

Appendix

A. Instructions

This is an experiment in the economics of decision-making. If you follow these simple instructions carefully and make good decisions, you could earn a considerable amount of money. The currency we will use throughout the instructions and the experiment is the Berkeley Buck. We will denote it as “$” and the exchange rate is $ 12 per US$ dollar. Please be aware that we do not expect any particular behavior from you or any other participant.

This session will be divided in 4 blocks, each comprising a series of situations in which you will have to make decisions. In what follows we will describe these blocks in chronological order as they will appear in your computer screen.

In block 1, you will face 4 situations. In each situation you will have to allocate a total of 10 tokens between yourself and another participant. The tokens may have different values in each situation. The other participant will be selected randomly at the end of the experiment. Also, neither you nor the other participant will receive information about each other's identity and about your decision in each situation. The computer will randomly select 1 of the 4 situations (with equal chance) to compute your payout based on your decisions. Symmetrically, at the end of the experiment you will be randomly matched to the decision made by another participant. Thus, you will have two ways of earning money from these situations. The first is from your allocation decision, and the second is from the allocation decision of another randomly matched participant.

Block 2 consists of one situation. You will have to send a message to another participant. This message will be available to the other randomly assigned participant at the end of the experiment. After he/she reads the message in his/her screen, he/she will make a choice. Your payout from this situation depends on the choice of this other participant.

Block 3 consists of one situation. You will have to allocate a sum of money between yourself and another participant. The other participant will be assigned randomly at the end of the experiment. In this situation your allocation may be reversed by the computer with some probability specified in the corresponding screen. Neither you nor the other participant will be told about each other's identity and about your decision. Thus, you will have two ways of earning money from this situation. The first is from your allocation decision (depending on whether is reversed or not), and the second is from the allocation decision of the other randomly matched participant. <STOP READING HERE>

In block 4 you will face 12 situations. Each situation will contain two types of scenarios: a non-interactive and an interactive scenario. In the non-interactive scenario, your decisions will be matched at the end of the experiment. In the interactive scenario your decisions will be matched immediately with a different participant in each round.

In the non-interactive scenario you will choose between playing two alternatives, A or B. The following screen shot shows an example:

[ADD SCREENSHOT OSPD HERE]

Each of these scenarios will feature different combinations of possible payoffs. In every scenario, you will be the Row player so your payoffs are the ones on the left of each cell. Both you and the other participant will have two possible choices. You can choose A or you can choose B. In this example:

- If you both choose A you will both get a payoff of $5.

- If you both choose B you will both get a payoff of $9.

- If you choose A, but the other participant chooses B, you will get a payoff of $15, but the other player will receive $4.

• Likewise, if you choose B, but the other participant chooses A, then you receive $4 and the other participant receives $15.

When choosing your move, you will not know the decision of the other participant. The other participant will not know your decision. For each scenario, your decision will be matched with the decision of another randomly selected participant at the end of the experiment. <STOP READING HERE>

After you decide between A and B, you will have to make a prediction about other participants' decisions. You will have to forecast how many of the other 23 participants in the experiment will choose either A or B in each scenario. You will be rewarded for the accuracy of your predictions. The formula to compute your reward is:

Maximum{0 , 8 – 1 * (Distance Between Your Prediction and Actual Decisions Other Participants)}

To explain the formula, we will focus on the prediction about A (because the prediction about B is 23 minus the Prediction about A). If your prediction coincides with the actual value you will get $8. An amount of $1 will be deducted for each unit above or below the actual number of participants who chose A (or equivalently B). Thus, if your prediction is more than 7 units away from the actual number of participants choosing A, you will receive $0. This calculation will be performed for each of the 12 scenarios. <STOP READING HERE>

From these 12 choices and predictions of each scenario, the computer will randomly select 6 to calculate the payouts. We emphasize that all the decisions described so far will be matched at the end of the experiment, so your actions will not affect the actions of other participants.

In the interactive scenarios, you will be randomly matched with a new participant in each of the 12 scenarios. We call each of these scenarios an Interaction. In each Interaction you will be matched with a different participant, so you will interact with one participant only once. Also, this participant will never be matched with any of the participants you will be matched with.

Each Interaction consists of two screens. In the first screen you will see the payoffs of the game and you may use the chat box on the left to communicate with the participant you are matched with in that Interaction. This screen will be shown for 30 seconds:

[ADD SCREENSHOT OF CHAT HERE]

In the second screen, you will have the opportunity to select your action. There will be no chat and you will have 30 seconds to make your decision. This is an example of a screen shot:

[ADD SCREENSHOT OF SIMULTANEOUS GAME HERE]

After each Interaction you will observe your and the other participants' choice (A or B) as well as your and other participant's payoffs. Your payoff in each Interaction will be based on your decision and the decision of the matched participant you interacted with. Your final payout will be calculated by the computer using 6 randomly selected Interactions (with equal chance) out of the 12. Recall that each situation in this block 4, comprising one non-interactive and one interactive scenario, will be repeated 12 times. <STOP READING

HERE>

Finally, at the end of the experiment, you will read the message a randomly matched participant sent you, and you will have to make a choice. You will also be given questionnaires that can yield some additional payoffs. After all questionnaires are completed, the computer will match the decisions of all the situations except for the Interactions (which were already matched in each of the 12 scenarios), and final payments will be made to each of you.

Each screen you see throughout the experiment has all the instructions necessary for the decision on that screen. Recall, during the session, all payoffs are expressed in terms of Berkeley Bucks. However, at the end of the session, all of your Berkeley Bucks will be converted at 12 Berkeley Bucks to 1 US$. Thus, in US dollars, your final payment will be between $5 and $30, depending on how you do.

Recall that at no time your true identity nor your final payout will be revealed to the other participants in this experiment.

Thank you very much and good luck!

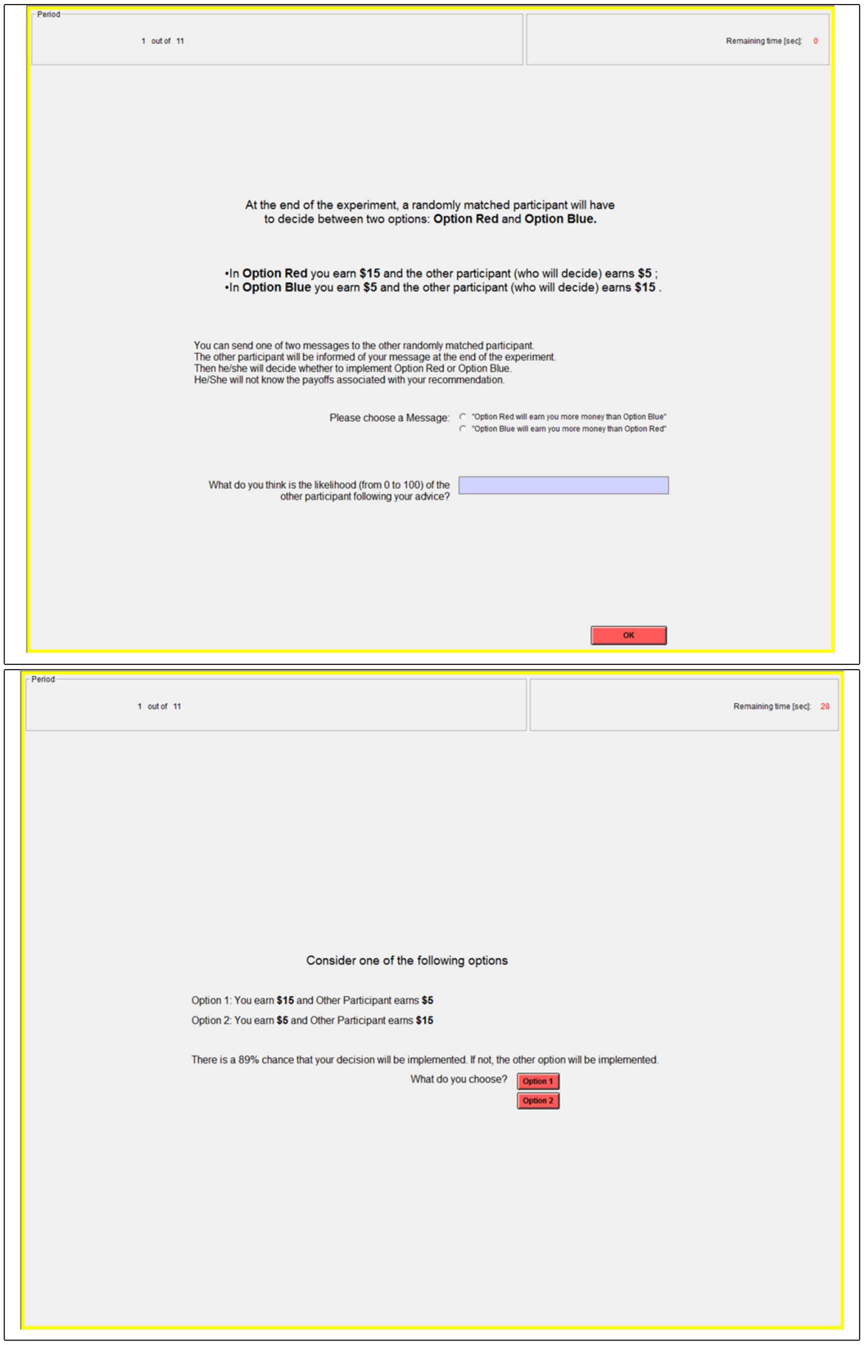
B. Screenshot Lying aversion elicitation

C. Treatment screen shots

D. Screenshots Risk Aversion test

E. Cooperation and own initiative

F. Cooperation and other's initiative



At the end of the experiment, a randomly matched participant will have to decide between two options: **Option Red** and **Option Blue.**

- In **Option Red** you earn **$15** and the other participant (who will decide) earns **$5** ;
- In **Option Blue** you earn **$5** and the other participant (who will decide) earns **$15** .

You can send one of two messages to the other randomly matched participant.
The other participant will be informed of your message at the end of the experiment.
Then he/she will decide whether to implement Option Red or Option Blue.
He/She will not know the payoffs associated with your recommendation.

Please choose a Message: ○ "Option Red will earn you more money than Option Blue"
○ "Option Blue will earn you more money than Option Red"

What do you think is the likelihood (from 0 to 100) of the other participant following your advice?

Consider one of the following options

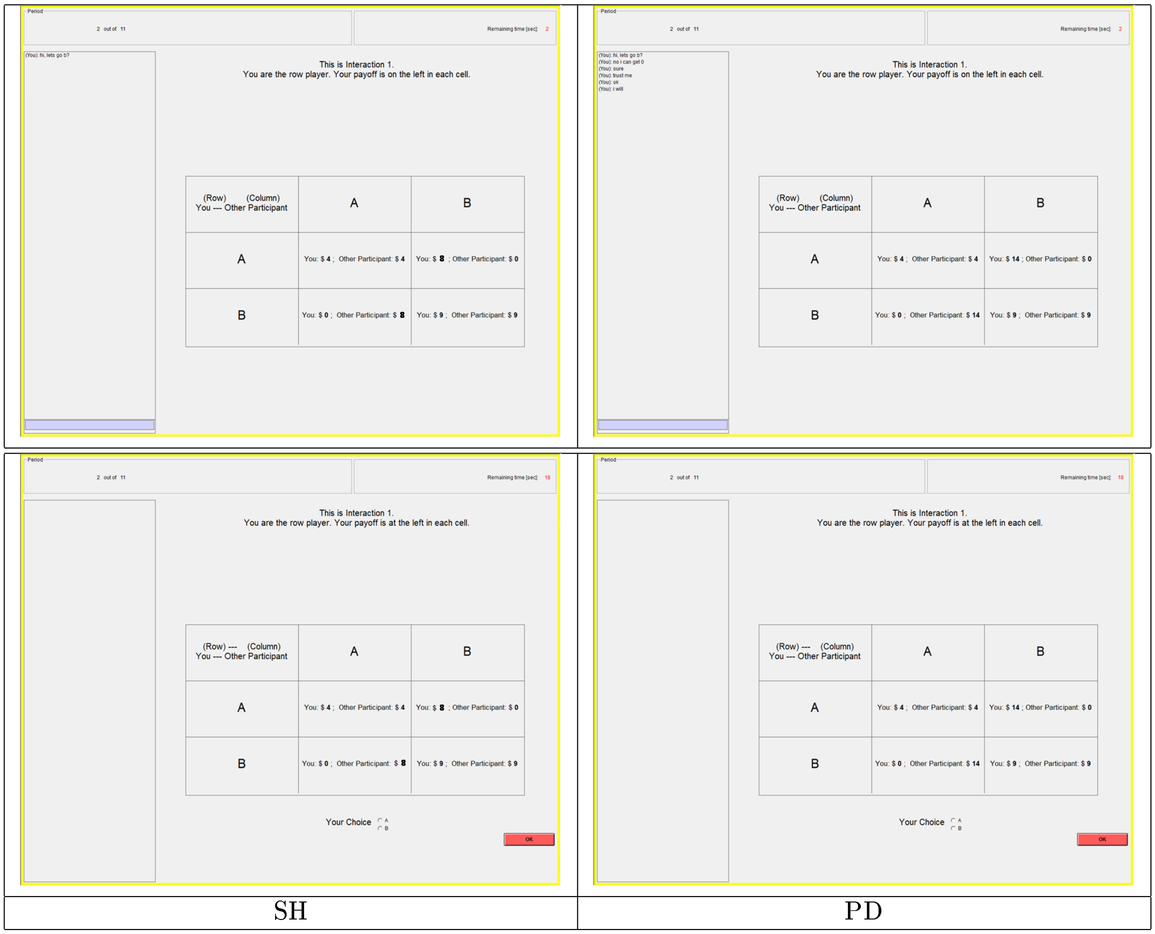
Option 1: You earn **$15** and Other Participant earns **$5**

Option 2: You earn **$5** and Other Participant earns **$15**

There is a 89% chance that your decision will be implemented. If not, the other option will be implemented.

What do you choose?

Figure 3: Screenshots lying aversion elicitation.



Period

2 out of 11

Remaining time [sec]: 2

(You): hi, lets go b?

This is Interaction 1.
You are the row player. Your payoff is on the left in each cell.

(Row) (Column) You --- Other Participant	A	B
A	You: $ 4 ; Other Participant: $ 4	You: $ 8 ; Other Participant: $ 0
B	You: $ 0 ; Other Participant: $ 8	You: $ 9 ; Other Participant: $ 9

Period

2 out of 11

Remaining time [sec]: 2

(You): hi, lets go b?
(You): no i can get 0
(You): sure
(You): trust me
(You): ok
(You): i will

This is Interaction 1.
You are the row player. Your payoff is on the left in each cell.

(Row) (Column) You --- Other Participant	A	B
A	You: $ 4 ; Other Participant: $ 4	You: $ 14 ; Other Participant: $ 0
B	You: $ 0 ; Other Participant: $ 14	You: $ 9 ; Other Participant: $ 9

Period

2 out of 11

Remaining time [sec]: 18

This is Interaction 1.
You are the row player. Your payoff is at the left in each cell.

(Row) --- (Column) You --- Other Participant	A	B
A	You: $ 4 ; Other Participant: $ 4	You: $ 8 ; Other Participant: $ 0
B	You: $ 0 ; Other Participant: $ 8	You: $ 9 ; Other Participant: $ 9

Your Choice ○ A ○ B

Period

2 out of 11

Remaining time [sec]: 18

This is Interaction 1.
You are the row player. Your payoff is at the left in each cell.

(Row) --- (Column) You --- Other Participant	A	B
A	You: $ 4 ; Other Participant: $ 4	You: $ 14 ; Other Participant: $ 0
B	You: $ 0 ; Other Participant: $ 14	You: $ 9 ; Other Participant: $ 9

Your Choice ○ A ○ B

Figure 4: Screen shots, treatments.

	(1) CG Pr$\{d_i$=C$\}$	(2) CG Pr$\{d_i$=C$\}$	(3) PD Pr$\{d_i$=C$\}$	(4) PD Pr$\{d_i$=C$\}$
Initiate	0.80***	0.72***	0.64***	0.68***
	(0.19)	(0.16)	(0.15)	(0.15)
Low Reciprocal-A. High Lying-A.	0.14	-0.55**	-0.13	0.15
	(0.42)	(0.27)	(0.37)	(0.46)
High Reciprocal-A. Low Lying-A.	0.09	-0.35	0.42	0.02
	(0.36)	(0.40)	(0.28)	(0.36)
High Reciprocal-A. High Lying-A.	0.18	-0.56	1.09***	1.33***
	(0.47)	(0.37)	(0.35)	(0.44)
Selfish		-0.57*		-0.80***
		(0.32)		(0.29)
InternalLocusofControl		-0.15***		0.13
		(0.06)		(0.09)
Extraversion		-0.46***		0.08
		(0.15)		(0.19)
Agreeableness		0.66**		0.70**
		(0.32)		(0.27)
Conscientiousness		-0.50		0.40*
		(0.35)		(0.24)
Neurotism		-0.40		0.15
		(0.26)		(0.22)
Openness		0.61***		-0.26
		(0.22)		(0.26)
ScoreCRT		0.04		-0.04
		(0.17)		(0.16)
RiskAversion		-0.57***		-0.11
		(0.18)		(0.14)
female		0.21		0.86
		(0.42)		(0.56)
Asian		-0.58		-0.73**
		(0.42)		(0.37)
White		-0.52		-0.43
		(0.45)		(0.54)
_cons	0.42	4.17	-0.95***	-4.91***
	(0.29)	(2.83)	(0.20)	(1.54)
N	576	540	576	540
pseudo R^2	0.075	0.315	0.130	0.274

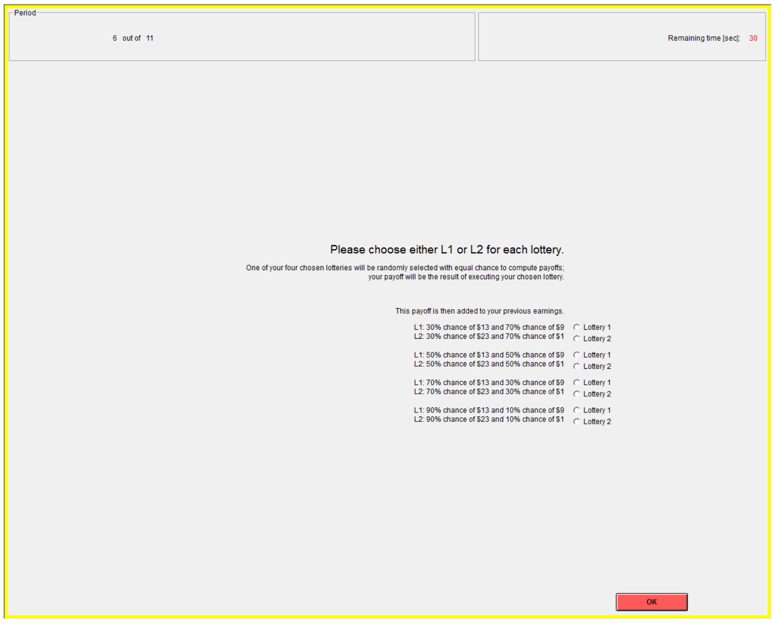
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Reduced form of cooperation on initiative, reciprocal altruism, lying-aversion and controls.

	(1)	(2)	(3)	(4)
	CG	CG	PD	PD
	Pr$\{d_i$=C$\}$	Pr$\{d_i$=C$\}$	Pr$\{d_i$=C$\}$	Pr$\{d_i$=C$\}$
Other Initiate	-0.05	0.19	0.30***	0.23**
	(0.15)	(0.16)	(0.10)	(0.11)
Low Reciprocal-A. High Lying-A.	0.08	-0.57**	-0.30	-0.09
	(0.42)	(0.28)	(0.38)	(0.45)
High Reciprocal-A. Low Lying-A.	0.05	-0.37	0.49*	0.19
	(0.37)	(0.43)	(0.27)	(0.35)
High Reciprocal-A. High Lying-A.	0.07	-0.60*	1.11***	1.46***
	(0.48)	(0.35)	(0.35)	(0.42)
Selfish		-0.60*		-0.69**
		(0.33)		(0.29)
InternalLocusofControl		-0.19***		0.11
		(0.06)		(0.09)
Extraversion		-0.41***		0.12
		(0.15)		(0.20)
Agreeableness		0.76**		0.55**
		(0.34)		(0.27)
Conscientiousness		-0.55		0.45*
		(0.38)		(0.24)
Neurotism		-0.39		0.21
		(0.28)		(0.23)
Openness		0.52**		-0.07
		(0.24)		(0.26)
ScoreCRT		0.02		-0.09
		(0.16)		(0.17)
RiskAversion		-0.66***		-0.11
		(0.17)		(0.15)
female		0.25		0.62
		(0.43)		(0.57)
Asian		-0.45		-0.69**
		(0.43)		(0.34)
White		-0.48		-0.58
		(0.46)		(0.52)
_cons	0.88***	4.88*	-0.82***	-5.03***
	(0.25)	(2.87)	(0.19)	(1.54)
N	576	540	576	540
pseudo R^2	0.001	0.278	0.099	0.242

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Reduced form of cooperation on other's initiative, reciprocal altruism, lying-aversion and controls.



Period

6 out of 11

Remaining time [sec]: 30

Please choose either L1 or L2 for each lottery.

One of your four chosen lotteries will be randomly selected with equal chance to compute payoffs;
your payoff will be the result of executing your chosen lottery.

This payoff is then added to your previous earnings.

L1: 30% chance of $13 and 70% chance of $9 ○ Lottery 1
L2: 30% chance of $23 and 70% chance of $1 ○ Lottery 2

L1: 50% chance of $13 and 50% chance of $9 ○ Lottery 1
L2: 50% chance of $23 and 50% chance of $1 ○ Lottery 2

L1: 70% chance of $13 and 30% chance of $9 ○ Lottery 1
L2: 70% chance of $23 and 30% chance of $1 ○ Lottery 2

L1: 90% chance of $13 and 10% chance of $9 ○ Lottery 1
L2: 90% chance of $23 and 10% chance of $1 ○ Lottery 2

Figure 5: Screen shot, risk-aversion test.