Taking the initiative. What motivates leaders?*

Lisa Bruttel† Urs Fischbacher‡

November 8, 2010

Abstract

Taking the initiative is a crucial element of leadership and an important asset for many jobs. We assess leadership in a game in which it emerges spontaneously since people have a non-obvious possibility to take the initiative. Combining this game with small experimental games and questionnaires, we investigate the motives and personality characteristics that entail leadership. We find efficiency concerns, generosity, and attention seeking as important determinants of leadership. Response time patterns and the results from the cognitive reflection test show that cognitive resources are relevant in the decision to lead.

Keywords: leading-by-example, social preferences, experiment

JEL-Classification: A13, C92, D03, D83

*We thank Kate Bendrick and participants of the ESA world meeting 2010 in Copenhagen and research seminars in Konstanz and Kreuzlingen for very helpful suggestions. Nathalie Hock and Natalie Zimmer provided valuable research assistance.

†University of Konstanz, Department of Economics, Box 131, 78457 Konstanz, Germany. E-mail: lisa.bruttel@uni-konstanz.de. Phone: 0049-7531-88-3214.

‡University of Konstanz, Department of Economics, Box 131, 78457 Konstanz, Germany. E-mail: urs.fischbacher@uni-konstanz.de. Phone: 0049-7531-88-2652. Thurgau Institute of Economics, Hauptstrasse 90, 8280 Kreuzlingen, Switzerland. Phone: 0041-71-677 05 12

1 Introduction

Imagine you share an apartment with roommates and the dirty dishes are piling up in the kitchen. With every additional plate, empty space in the kitchen disappears. Eventually, the filth tolerance level of one of the roommates is surpassed and this poor person starts doing the dishes. Often this initiative encourages roommates to follow the good example. Everybody likes having such leaders as roommates, and human resource departments have a strong interest in recruiting these personalities as well. Therefore, it would be desirable to know the determinants of this behavior. What motivates people to take the initiative and set the good example? Taking the initative by giving a good example is a particularly interesting type of leadership because it is voluntary from the side of the followers as well as from the side of the leader. In this respect it differs from concepts of leadership that are often discussed in the literature (Yukl 2009; Kouzes and Posner 2007) where the focus is on advising designated leaders. The notion most closely related to our research topic is charismatic or transformational leadership (Bass and Avolio 1994).

Taking the initiative is not only important as a leader selection device. In any bad state of social interaction it is desirable that someone breaks the vicious circle by giving a good example, but only some actually do so. We are interested in the characteristics of those who do. In a discussion with a teacher, who raises at first a point about an unfair exam? Who starts with using nets with larger mesh size at a lake that tends to be overfished? Another economically relevant example is the open source community. Without people like Linus Torwalds who initiate a project and contribute a significant code base, open source projects will never start (Lerner and Tirole 2002).

In this paper we aim to understand the motivations of voluntary leaders to set the good example. As we have seen in the illustrations above, leading-by-example is a behavior in which a leader aims to induce others to act in the groups' joint interest by demonstrating the behavior that they should imitate. The situation is characterized by the following features. First, the leader has to understand and anticipate the response of the followers. The action that motivates others to follow is not always easy to find. Second, leading-by-example is initially costly (i.e. the leader has to incur some cost or forgo possible gains, which might be recovered later) but successful leadership is beneficial for the group. Concerning the characteristics of leaders, we therefore expect cognitive skills to be important for leading as well as the willingness to use these resources. Further, leading-by-example could be driven by selfish profit maximization if the leader believes that leadership pays in the long run. However, the benefit accrues in the future, which implies that patience could play a role in peoples' willingness to lead. Risk preferences are also expected to matter for leadership behavior because whether it pays off is not certain. Also other-regarding preferences should matter, because when the followers contribute less than the leader, the leader's earnings could be lower than that of the followers. Nevertheless, the group as a whole will benefit. Thus, motives like altruism, generosity, efficiency and envy are likely determinants of leadership. Finally, leaders get a special role by their behavior. Thus, they get a positive image, which could also be a motivation to lead.

Because of the incentive structure of leading-by-example, this kind of leadership is often studied by introducing a sequential move structure in public good experiments (Gächter and Renner 2006; Güth et al. 2007; Gächter et al. forthcoming; Moxnes and van der Heijden 2007; Pogrebna et al. 2009; Potters et al. 2007; Levati et al. 2007). These studies focus on the mechanism of leadership and typically show that groups with leaders on average contribute more than groups without, but only due to the higher contributions of the leaders. Arbak and Villeval (2007) investigated the motivations of leaders in this situation and combined different variants of a two-stage public good game with personality tests. They showed that social concerns are a driving force for at least some of their leaders.

Public good games capture nicely the incentive structure of leading-by-example. However, in these games it is obvious to all players what constitutes the good example and, therefore, they do not cover the innovative facet of the problem. Furthermore, it is always clear to the subjects that the experiment they are participating in is about leading and following. This may induce experimenter demand effects (Zizzo 2010) possibly manipulating leadership in either direction. It may reduce leadership because involuntary leaders perform worse; or it may enhance leadership, because even natural non-leaders infer from the experimental design that leadership is socially desirable. In order to avoid the experimenter demand effect problem and to include the innovative element of leadership into the design, we use a completely different setup to study leadership. In our design, there is no predefined leader. So we cannot only address the question of whether a person accepts to be a leader when she is assigned the role but also whether a person chooses to voluntarily lead. We build on a game introduced by Dufwenberg and Gneezy (2000) to study the Bertrand paradox and interpreted by Bruttel (2009) in terms of leading-by-example. In this two player game, people choose a number between 2 and 100.

The person setting the lower number gets a corresponding payoff; the other gets nothing. In the case of a tie both receive half the price. This game is repeatedly played with changing pair composition, but all eight subjects in a group are informed about the decision of all players in their group after the end of each round. In the experiment, the numbers decrease to a rather low level and earnings are low. Eventually though, one player raises the number to a very high level. Many of the other players follow, so a temporary increase of the numbers occurs. In the present study we consider subjects which initiate a number increase as leaders. Different from other experiments about leadership, the leading behavior in this experiment is neither explicitly nor implicitly induced by the experimenter. There is no explicit assignment of the leading role to a certain subject. Decision making occurs simultaneously, so no player has a distinct role. Nevertheless, this design has proven to produce reliable rates of leadership. We use the sudden number increase in this game to identify natural leaders and connect this classification to the decisions in other games and questionnaires eliciting other-regarding preferences, beliefs, risk attitude, cognitive abilities, and other personality characteristics.

According to our results, leaders are characterized by above-average cognitive skills and predominantly male. They have strong preferences for efficiency, generosity, and against advantageous inequality, and do not primarily seek to maximize their personal monetary benefit. Instead, leaders are motivated by the positive public image of being a leader. They have accurate beliefs about the extent to which others will follow their example but rather underestimate the probability that other players will act as leaders as well. We do not find an impact of personality traits or risk attitude on leading-by-example.

In section 2 of this paper, we present the design we use in the experiment. In section 3 we provide the behavioral predictions. Section 4 presents the main findings and section 5 concludes.

2 Design and Procedures

2.1 Design

In order to determine the motivations of leaders, we combine a game in which spontaneous leadership occurs regularly with a series of experiments and questionnaires that allow to measure beliefs, risk and social preferences and other motivations. We start with the explanation of the game that we use as our leadership game.

The basic design of the game in our experiment is a variant of the stylized Bertrand pricing game in Dufwenberg and Gneezy (2000). In this game, two participants simultaneously choose a number from the interval [2,100]. The participant choosing the lower number wins the game. The prize is equal to the winning number. In case of a tie, each participant gets half of the prize. The game is repeated for 30 rounds. Players are divided into groups of eight participants. In each round, the eight participants in one group are randomly matched in pairs of two. Thus, four pairs play the game simultaneously in a group of eight participants.

It was known from Bruttel (2009) that the behavior in the experiment crucially depends on the feedback condition. If subjects only get feedback from their own pair, then the chosen numbers continuously decline. However, if subjects are informed about all decisions from their group, typically a cyclical movement of average numbers is observed. In the beginning, the eight numbers within a group are uniformly drawn from the set between 2 and 100. During the first few rounds, average numbers decrease, because all participants try to choose a number slightly lower than most of the others. Strategies in this phase of the game are well described by directional best reply to the distribution of numbers in the previous round. After some rounds, the group reaches a relatively low level of numbers. Eventually, one of the players chooses a very high number (often even 100), presumably to signal the others that they should coordinate at a higher level. Such signaling is only possible because of the group feedback after each round. This behavior is in no way induced by the experimental instructions, but appears to be very robust. We consider the initiative to coordinate at a higher level to be endogenous leading-by-example.

In our experiment, after each round, the subjects were informed about the decisions in their group. Furthermore, all 8 numbers were made publicly known in the group, ordered by size of the number. Thus, players not only received feedback about their own number and the number of their partner in this period, but also about the other participants whom they did not meet in the current round but might meet in the next rounds.

Each number choice decision in this main part of the experiment was surrounded by a belief formation stage and a publicity choice stage. In the belief formation stage before the number choice, players had to submit beliefs about the minimum, maximum, and average number of the other seven players in the next round. For each of the three values, they had to submit a probability distribution over the intervals 2-20, 21-40, 41-60, 61-80, 81-100. To facilitate submission of their beliefs, they were provided a graphical tool on the computer screen. Figure 8 in the Appendix shows a screen shot. The bars of the single intervals could be moved with mouse clicks. A click on "update" next to one of the distributions automatically increased or decreased all five bars proportionally to balance the sum of weights to 100 percentage points. If participants were done with their belief formation, they had to click "next". In the beginning of the next round, their past estimates were shown as default values and could be adapted with the same procedure. The quality of their prediction for each of the three values was determined with the quadratic scoring rule (Brier, 1950). They received a payment proportional to this measure.

By asking subjects before their decision about their belief about the probability distribution of the maximum number of the other group members in the next period, we learn how likely they think it is that someone else will lead. In particular, we need the probability weight leaders assign to the categories equal to or larger than their own leading number. The estimated average number of the other group members one round after a leading number provides an approximation of the leader's belief to what extent the others will follow. For our analysis, we re-calculate the estimated average from the submitted probability distribution. By comparing the beliefs of leaders to the beliefs of other not leading participants, we learn whether leaders are different from others with respect to their estimate of the benefits of leadership. The stronger the increase in the average number after a leading number, the higher are the potential gains to a participant undercutting opponents by a small amount. If leaders systematically overestimate others' average numbers after a leading number, this would indicate that leaders lead because they overestimate their monetary benefits from leading.

In addition to the maximum and the average number, we asked players to submit their belief about the minimum number of the other group members which we do not need for the analysis at all. We elicit beliefs in such a detailed way to receive an accurate belief of leaders on whether there will be another leader. Asking for this probability of the maximum number only, however, could introduce the experimenter demand effects again which we were able to avoid by the design of the main part. In order not to lead subjects into thinking about leadership, we therefore included the minimum belief and applied the distributional belief elicitation procedure to all three values, minimum, maximum, and average.

In the publicity stage after the number choice, we allowed players in the given round to give up anonymity and publish their seat number on the other participants' computer screens beside their own chosen number. Publication of the seat number in one round cost 10 points and could be decided upon by ticking a box on a separate screen after the number choice. Use of this feature allows us to control for whether appreciation by others motivated extraordinary number choices.

Treatment	Description
1	Distribution games (efficiency, inequality, generosity)
2	Risk elicitation
3	Belief trial phase
4	**Number choice game**
5	Feedback about outcomes and payoffs
6	Strategy questionnaire
7	Time preference questionnaire
8	Leadership questionnaire
9	Cognitive reflection test
10	Further questionnaires

Table 1: Order of Treatments

Before the main part of the experiment, we conducted some short games to elicit preferences for efficiency and generosity, inequality aversion and risk attitude. Table 1 includes an overview of the order of the different games in the experiment. We applied the same order of these experiment to all the subjects. This procedure has the disadvantage that there might be spillovers between the games for which we do not control. However, we are interested in the difference between leaders and non-leaders and as long as there is no interation between the type and the spillover, we can draw valid conclusions. Furthermore, a fixed order has the advantage to avoid introducing extra noise into the data. The first part of the experiment was a series of seven simple two-player distribution games using the strategy method, similar to Engelmann and Strobel (2004). As Bruttel (2009) argues, there seem to be spillover wealth effects from the main experiment to the decisions in such distribution games. For this reason, we conducted these games before the main part of the experiment and not afterwards. In each game, participants had to choose between two distributions of money between themselves and another player. Table 2 shows the options between which player 1 could choose and their payoff consequences. The seven games were designed in order to create tradeoffs between selfishness, equality and efficiency. In the first column, there is a tradeoff between selfishness and equality on the one hand and efficiency on the other. In particular, envious people will go for option A. The second column contains games with a tradeoff between selfishness on the one hand and equality and efficiency on the other hand. In the third column there is one game. In this game there is a tradeoff between equality in the form of envy and efficiency. The roles of players 1 and 2 were randomly assigned to the players after they had decided for both roles. Only one out of the seven games was randomly selected for payment. After completion of the seven choices, we elicited risk attitude using the Holt and Laury (2002) procedure. The random draws from this part of the experiment and the corresponding payoffs were revealed only after the main part of the experiment.

	Equality vs. Efficiency		Selfishness vs. Equality and Efficiency		Envy vs. Generosity	
	Option A	Option B	Option A	Option B	Option A	Option B
Player 1	2	1	5	4	1	1
Player 2	2	4	2	4	1	3
Player 1	2	1	6	5		
Player 2	2	5	2	5		
Player 1	2	1	7	6		
Player 2	2	6	2	6		

Table 2: Parameters in the distribution games

2.2 Procedures

The experiment was computerized using z-Tree (Fischbacher 2007). A total of 96 students, 44 males and 52 females, from various disciplines took part in the experiment, divided into 12 groups of 8 participants each. They were recruited via ORSEE (Greiner 2004). The experiment took place in the *Lakelab*, the laboratory for experimental economics at the University of Konstanz between December 2009 and June 2010. Sessions lasted between 2 and 2.5 hours. The experimental currency was points, with 30 points converted into 1 Euro (between $1.20

and \$1.50 at the time of the experiment) after the experiment. On average, participants earned 28.77 Euros in the experiment. The protocol during the experiment was as follows: After welcoming participants and explaining the main rules for participation in the experiment, they were randomly assigned seats in the laboratory. At their place, they read short general instructions about the sequence of experiments they would participate in. For the first two parts of the experiment, subjects received instructions on their computer screen and made their decisions immediately after reading the instructions. For the main part of the experiment, they received written instructions explaining decision-making and its consequences as well as the belief formation stage including the payment method with the quadratic scoring rule and regarding the publicity choice stage. Next they were given the possibility to familiarize themselves with the computer screen for the belief formation. Then the experiment started. At the end of the session, the participants were asked to complete several questionnaires. Players first had to answer a questionnaire about their decisions in the number choice game. After that, they were asked to fill in several questionnaires, including the big five questionnaire of Brandstätter (1988), a locus of control questionnaire, a cognitive reflection test (Frederick 2005), and a socioeconomic questionnaire.

3 Behavioral predictions

In this section, we focus on our research question - what motivates leaders. At the beginning of the next section, we will give the exact description of how we classify leaders. For now, we just note that if there is common knowledge about rationality and selfishness, subjects should choose 2 as their number. So, even when subjects try to coordinate on higher number at the beginning, directing the behavior towards the best reply of the previous period will cause a decline in the numbers and, hence, in the payoff (Selten and Stöcker 1986). A subjects displays leadership when she breaks out of this vicious circle and increases her number. In this section, we discuss the potential motivations for this behavior.

First, let us consider the selfish motivation to lead. Some leaders in our experiment might initiate a number increase not for the purpose of the benefit of the group, but rather because they intend to undercut others at a higher level in the next round. Such selfishly motivated leadership crucially hinges on the belief to which extend the other players will follow. Actually, Gächter et al. (forthcoming) find that cooperative leaders have over-optimistic beliefs about the cooperativeness of followers, and that this can (aside from social motivations of leaders) explain their high contribution as first mover in a sequential public good game. These over-optimistic beliefs might be a consequence of the false consensus effect (Ross et al. 1977). It seems likely that such over-optimism is not only present for the randomly assigned leaders in the sequential public good game in Gächter et al. (forthcoming) but also in the context of our number choice game. Thus, our first prediction is:

Hypothesis 1 *Leaders over-estimate the average numbers of their group members after leading.*

Our second hypothesis refers to other-regarding preferences. Successful leadership will provide a higher payoff for the group, but it is potentially costly for the leader and it could in particular create inequality that is disadvantageous for the leader. In the framework of a sequential public good game Arbak and Villeval (2007) find that voluntary leadership is related to preferences for efficiency and generosity (measured as charity donation behavior). In Table 2, option B is always the efficient outcome. If leaders care more about efficiency than non-leaders, we expect that leaders are more frequently choosing option B. In the first column, efficiency does not only go against selfishness, it creates also disadvantageous inequality, which envious people will dislike. Since leaders risk disadvantageous inequality, we expect them to be more tolerant towards disadvantageous inequality and to choose option B more frequently that non-leaders in particular in the games in the first column of Table 2.

Hypothesis 2 *Leaders have stronger pro-social attitudes than non-leaders. They attach a higher value to efficiency, they are more generous and they are more willing to accept disadvantageous inequality.*

Leadership could be a signal of prosociality. As a prosocial attitude is generally seen as a positive trait it might be that leaders lead because they want to signal their "good character" to others. In our experiment, we offer participants an opportunity to make their seat number publicly know. We expect that leaders use this option more frequently than non-leaders.

Hypothesis 3 *Leaders are more likely to give up anonymity than non-leaders. Leaders are more likely to give up anonymity in their leading rounds than in other rounds.*

Leadership is risky. When deciding to provide a good example, the leader hopes that the others will follow the example. In this case, future social welfare and potentially the leader's individual payoffs will increase. However, the leader cannot be sure that others will in fact follow. A risk averse player might therefore be reluctant to lead even if this person was willing to set the example if he was guaranteed that others will follow. We derive our next hypothesis:

Hypothesis 4 *Leaders are less risk averse than non-leaders.*

The considerations of a leader before deciding to lead are relatively complex. In the beginning of the game most players, including the later leaders, follow a best reply strategy against the distribution of numbers in their group in the previous round. This best reply dynamic leads to decreasing numbers from round to round, because all players try to choose a lower number than their representative opponent. The leader first has to understand this dynamic that all players react in a similar way to the group feedback. Second, the leader must be innovative in exploiting this behavior. By drastically increasing the own number, the leader manipulates the distribution of numbers the others are reacting to and reaches a temporary coordination of the group at a higher level of numbers. This understanding and manipulation of the dynamic decisions in this game requires a lot of innovation, creativity and cognitive ability. It requires also the willingness to break out of the simple responding to the other players' behavior. The cognitive reflection test described in Frederick (2005) captures the essence of these abilities. This is summarized in the next hypothesis.

Hypothesis 5 *Leaders have a higher score in a cognitive reflection test than non-leaders.*

According to Frederick (2005), the score in a cognitive reflection test is on average higher for males than for females. Furthermore, Arbak and Villeval (2007) hypothesize that particularly male participants may be concerned with maintaining a positive public image as men in their sample act more often as voluntary leaders than women. The latter result is also found in Gächter et al (forthcoming), though it is not significant there. Matched with the information about participants' gender, we can also test whether male leaders are more publicity seeking than female leaders. We formulate this as our next hypothesis.

Hypothesis 6 *Men are more likely to lead than women. In particular, men are more likely than women to give up anonymity as leaders.*

Leading-by-example is an optimistic act to improve the inefficient situation the group is in after a phase of mutual underbidding. Taking the initiative, the leader has to trust in his ability to change the circumstances of the interaction. We therefore hypothesize that leaders have an internal locus of control. With respect to the other personality traits we have no specific hypotheses concerning their impact on leadership.

Hypothesis 7 *Leaders have an internal locus of control. Other personality traits do not have a significant impact on leadership.*

The previous hypotheses were all about behavior. We conclude with a prediction concerning response time (for an interesting application or response time to economic decision making see Rubinstein 2007). Given the cognitive effort necessary before a player decides to lead, we expect that leading decisions take longer than standard decisions, when players try to maximize their expected profits against the distribution of numbers within their group. We furthermore expect that the decision to lead develops over several rounds. This implies that reaction times will slow down one or two rounds before the leading decision. Regarding the non-leaders, we expect a similar pattern. Their decisions may slow down after a leading number of someone else, because following a leader implies a new consideration between best reply behavior and following the leader's example.

Hypothesis 8 *Leading decisions are slower than decisions of currently not leading leaders. Decision times of leaders slow down more than one round before the leading number choice.*

4 Results

We start the review of our results with an overview of the average numbers in all 12 groups. Figure 1 illustrates them. In all groups, average numbers fluctuate quite considerably, indicating dynamics within the groups. Average winnings numbers follow a very similar pattern. Looking at the initial phase of the game, we see that average numbers in most groups decrease from round to round, while in some groups (groups 2, 5 and 9) they start by increasing. In these three groups, at least one player chooses the number 100 in the first round which triggers the first upward movement of average numbers right in the beginning of the game. Thus, the number 100 in the first round of the game already seems to be an instrument of leadership.

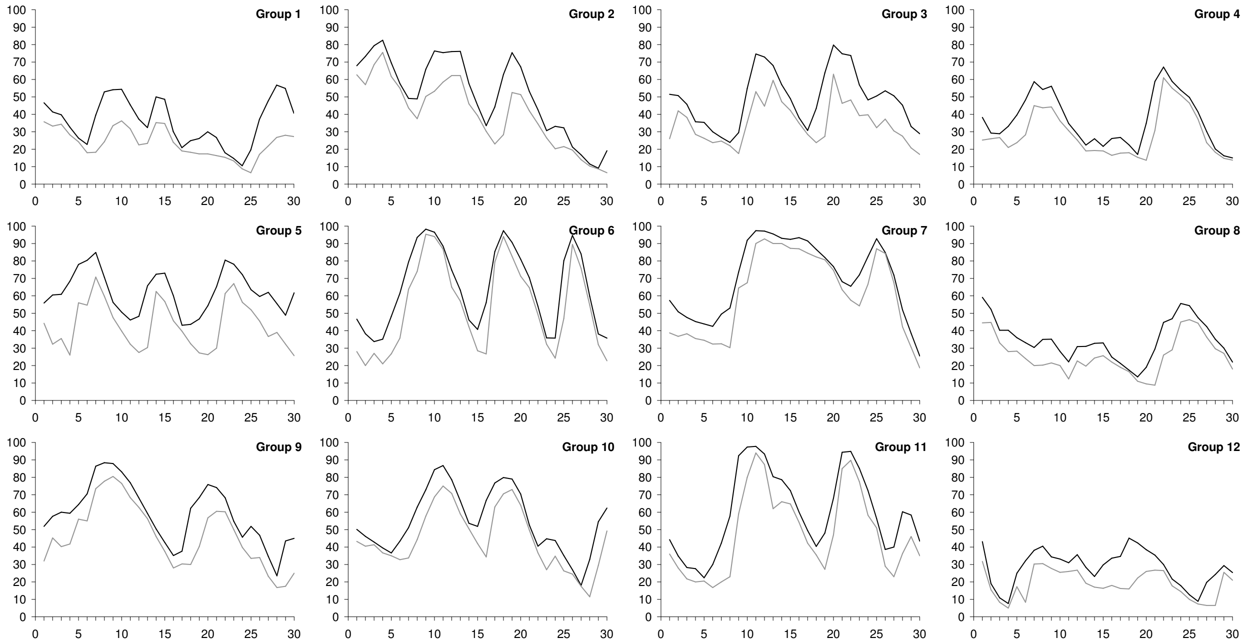

Figure 1: Average numbers (black line) and average winning numbers (grey line) in the 12 groups over all 30 rounds.

Let us next consider groups with decreasing average numbers in the first rounds. In almost all of these groups the downward trend of average numbers stops after at most 10 rounds and turns into an increase instead.[1] This later increase is always initiated by one player (sometimes also two at the same time) increasing the number substantially. Different from leadership in the first round, these leaders in later rounds do not necessarily increase their number to 100. This confronts us with the problem of disentangling intended leadership from casual number increases without a leading purpose. In order to identify intentional leaders systematically, we use a refinement of the definition introduced by Bruttel (2009). There, a leading number has to be more than 30 points larger than the leader's number in the round before and it has to be larger than all numbers of all other players in this group in the previous round.[2] We base our decision also on these criteria. In our understanding leadership contains the expectation that others will follow. We therefore define that leadership can only occur until two rounds before the end of the game. Later high numbers may be observed for others reasons but cannot be motivated by the intention to lead.

[1]An exception is group 8, in which average numbers continue to move downward for more than half of the game. In this group, two players lead by choosing 100 in the first round but give up after the others did not follow immediately. After this experience, it took a relatively long time until some else initiated the final increase of average numbers.

[2]In Bruttel (2009) no leading numbers in the first round of the game were considered.

Definition: A number n_{it} of player i in round t is called *"leading number"* if one of the following conditions is satisfied:

1. $n_{it} = 100$ if $t = 1$ or
2. $n_{it}(t) > n_{j,t-1} \forall j \in [1; 8]$ and $n_{it}(t) > n_{i,t-1} + 30$ if $t \in [2; 28]$

The player i who places the leading number is called a *"leader"*.

With this definition, we identify 39 out of our 96 subjects as leaders. Accordingly, 57 subjects are classified as non-leaders. Once leaders are identified according to the above definition we group them into the two subcategories "early" and "late" leaders. Early leaders are the first leaders in their group, late leaders are all subsequent leaders.[3] Late leadership is a weaker variant of leading for several reasons. The innovative aspect of leadership disappears if the leader has already observed someone else leading. Thus, late leaders do not necessarily have to have above-average cognitive skills. Late leaders have also already observed the reaction of their group to leadership. Therefore, they have an easier task in forming a belief about the potential gains and losses of leadership for the leader and the consequences for group efficiency. Out of the total 39 leaders 16 are early leaders[4] and 23 late leaders. In each group, we are able to identify at least one early and one late leader. The only exception is group 7, where we have two early leaders in the first round, but no late leader. More than half of the leading numbers had the value 100.

4.1 Does leadership pay for the leader - and what do they expect?

We start our analysis of the characteristics of leaders with the question of whether leadership is profitable for the leaders. Figure 2 depicts average profits of early, late and non-leaders. It

[3]We excluded the data from one additional matching group from the analysis. In this group, one participant continuously set the number 100 over almost the whole duration of the game, between round 5 and round 27. This disabled us from classifying the remaining seven participants in this group into late leaders and non-leaders, because they had no chance to lead during the whole experiment, even if they wanted to. Therefore, we decided not to consider this group at all.

[4]There are more early leaders than matching groups, because it happened four times that two subjects led early in the same round.

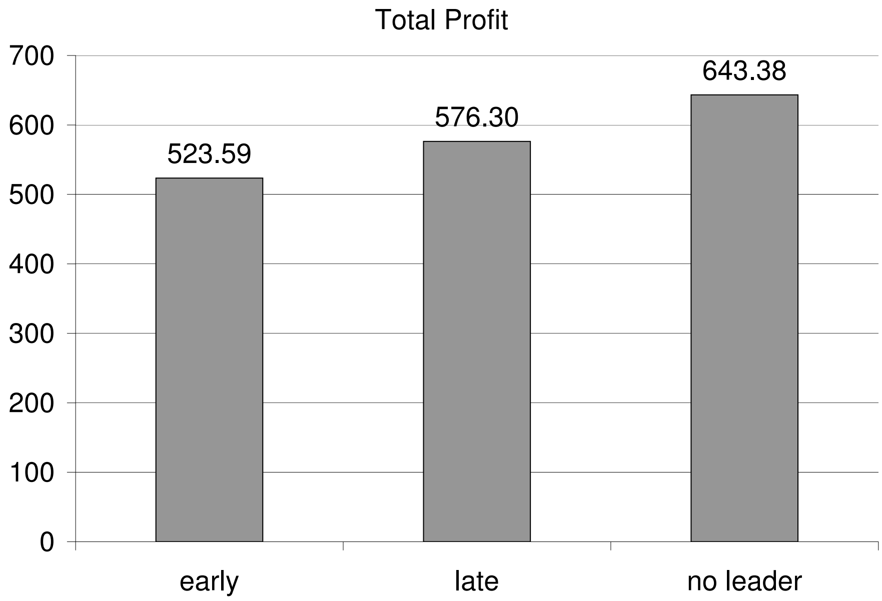

Figure 2: Profits in the number choice game.

shows that leaders earn significantly less than non-leaders (p-value = 0.0121).[5] The difference is particularly strong for early leaders and non-leaders (p-value = 0.0096). This finding is not surprising given that leaders deliberately forgo the possibility to win their match while leading. More relevant for the motivation of leaders (and much more difficult to answer) is the question whether leaders benefit from leading compared to the counterfactual situation in which they do not lead - and whether they correctly anticipate their net monetary loss or benefit from leading. We cannot answer the first question because we do not have a reference point which we could compare leaders' profits with. However, we can say that they are very good in anticipating the extent to which others will follow their good example, which is the basis for their own expected loss or gain from leading. We use the beliefs submitted for the average number of the seven fellow participants. If leaders overestimated the average number of the others, it was very likely that they overestimated their gain from leading as well. Leaders are generally very good in estimating the reaction of others after their leading bid. If at all, they slightly under-

[5] For a statistical comparison of leaders and non-leaders we treat each matching group of eight participants as one independent observation. Thus, we consider 12 independent observations. Within each group, we average the scores for each measure, e.g. the profit, over all early leaders, late leaders and non-leaders separately. All reported significance levels in this paper are then obtained (if nothing else is stated) in two-sided Wilcoxon signed rank tests testing the measures of (early or late) leaders against the non-leaders in each matching group. We do not correct for multiple hypothesis testing since we have ex ante hypotheses for almost all tests that we conduct.

rather than overestimate the average number of the seven other participants in the round after their leading bid (by about 4 units on the scale from 2 to 100). The quality of their estimate does not depend on whether they are currently leading or not, and it is also not different from the quality of the estimates of the non-leaders. We conclude that over-optimistic beliefs as in Gächter et al. (forthcoming) are not driving leadership in the framework of our number choice game. This makes it unlikely that selfish motives are the major driving force for taking the initiative.

Result 1 *Leadership does not pay, compared with the income of non-leaders.*

Result 2 *Leaders have realistic beliefs about how much the followers respond to their leading decision.*

The decision to lead might not only depend on the beliefs whether others will follow but also on the belief whether someone else in the group will take the initiative instead. To capture this belief, players had to submit an estimate for the probability distribution over the intervals 2-20, 21-40, 41-60, 61-80, 81-100 of the maximum of the other seven numbers in their group in each round. Figure 3 illustrates the actual average values of the others' maximum and the corresponding beliefs of leaders who are currently leading, leaders who are currently not leading and non-leaders. The data behind this illustration contains only values from rounds where leading was generally possible, i.e. rounds in which the maximum number in the round before was smaller than 100 and the minimum number was smaller than 70. To approximate the belief of whether there is another leader, we usually consider the percentage weight given to the category that the maximum of the other seven numbers falls into the interval from 81 to 100. If the leading bid was smaller than 100, we sum up the percentage values of the intervals around and above the leading bid. A similar procedure was used to determine whether there actually was another leader.[6] In line with the argumentation of a false consensus effect (Ross et al., 1977) leaders would generally overestimate others' willingness to lead (p-value = 0.2237) while non-leaders underestimate the probability that there would be a leader (p-value = 0.0022). However, in their leading round, leaders strongly (p-value = 0.0499) underestimate others' willingness to lead. In our interpretation, this pessimistic belief about the probability that others will lead additionally motivates leaders to take the initiative.

[6] Here, we used a different leader definition than above. Basically all players are treated as alternative leaders who choose a number that would be a leading number if it was the number of the actual leader.

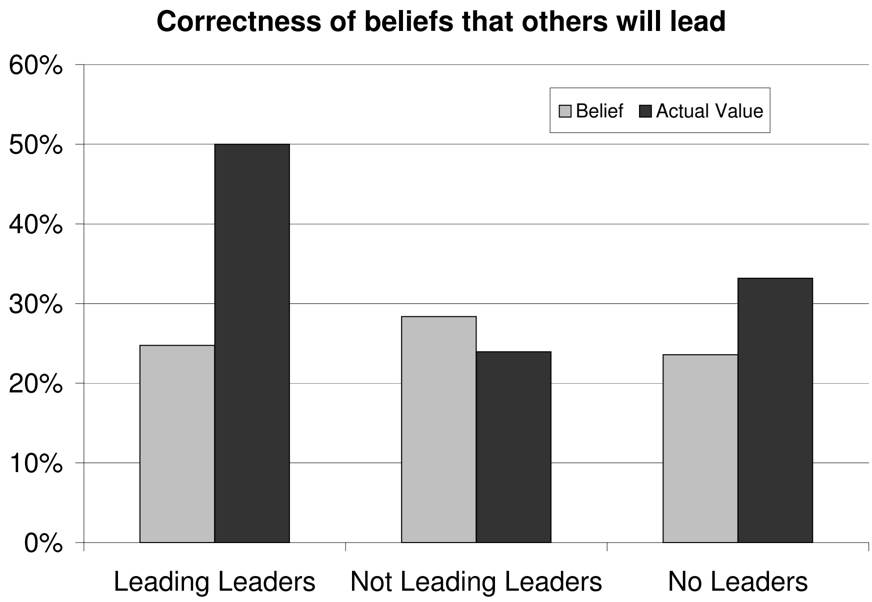

Figure 3: Belief that others will lead.

4.2 Leaders attach a high value to efficiency

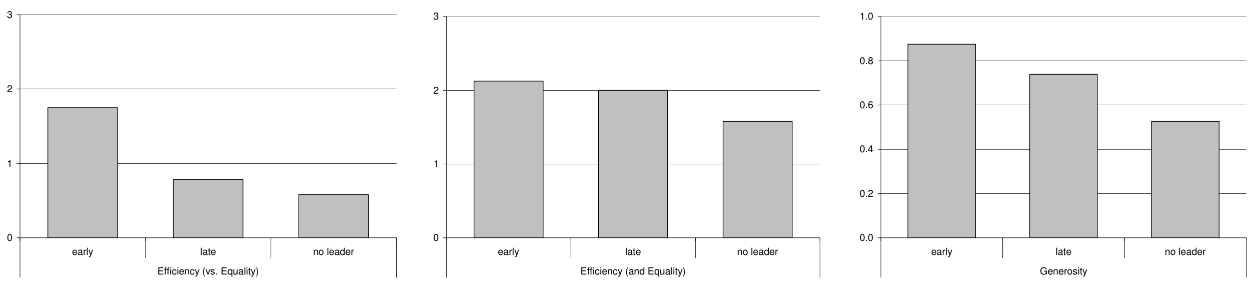

Figure 4: Valuation of efficiency, inequality and generosity.

Figure 4 summarizes the decisions of all participants in the role of player 1 in the distribution games. Leaders have stronger preferences for efficiency and they are more generous. We find the most notable difference for early leaders when there is a conflict between efficiency and equality. In this game, early leaders have an efficiency score of in almost 2 of 3, while the late leaders and the followers chose the efficient option in less than one case. Table 3 summarizes the significance levels. For efficiency concerns and generosity it holds that early leaders choose the efficient or generous option more often than all other players, and late leaders score higher than non-leaders. With respect to inequality aversion, the major difference is between early leaders and non-leaders, with no significant distinction between early and late leaders or between late leaders and non-leaders. Taken together, the results from the distribution games indicate that concerns for others' outcomes are a driving force for leading-by-example. Leaders are more pro-socially minded than non-leading players.

Risk aversion, as measured by the Holt and Laury (2002) lottery procedure has no significant effect on leadership. Early leaders are a little less (average number of safe choices 5.31), late leaders a little more (5.70) risk averse than non-leaders (5.46), but the p-values in Table 3 are far from any reasonable level of significance. This contradicts our hypothesis that leaders have a more positive attitude towards risk than non-leaders. The reason might be that the risk of leadership is different from (and hardly correlated with) the risk measured with the Holt and Laury (2002) lotteries. Their procedure generates risk as random draws between lotteries while the risk of leadership is a behavioral risk depending on the reaction of followers. The former requires calculation of expected values while the latter depends on the ability to predict others' behavior.

	early vs. no leader	early vs. late	late vs. no leader
Efficiency	0.0186	0.0973	0.2662
Inequality Aversion	0.0593	0.7310	0.0408
Generosity	0.0307	0.1297	0.0748
Risk aversion	0.7527	0.3936	0.8937
Male	0.0223	0.0152	0.4485

Table 3: Statistical tests for decisions in the distribution games, risk attitude, and gender.

Result 3 *Leaders attach a high value to the maximization of others welfare. Early leaders are more willing to accept disadvantageous inequity than others when it is in conflict with efficiency.*

4.3 Leaders like attention

Leaders might lead not because they want to do something good for their group, but because they want their group to see that they are doing something good. To test for the motivation of a positive public image our experiment introduced a publicity feature. In each period, players can pay to give up anonymity. They have to click a button on a separate screen after their number choice and have to pay 10 points. If they choose to give up anonymity, their seat number is displayed next to their selected number on the feedback screen for all players in the group in that round. Figure 5 shows that leaders indeed like to be noticed. In leading rounds, around 16 percent of the leaders opt for publicity; one round later (when the leaders often still have the highest number in their group), this holds for 13 percent of them. In comparison, non-leaders pay for publicity in less than 1 percent of their decisions and currently not leading leaders in only 3 percent. The differences between leaders and non-leaders as well as between leaders in t or $t-1$ and currently not leading leaders are statistically significant (p-values equal to 0.0043 and 0.0958, respectively).

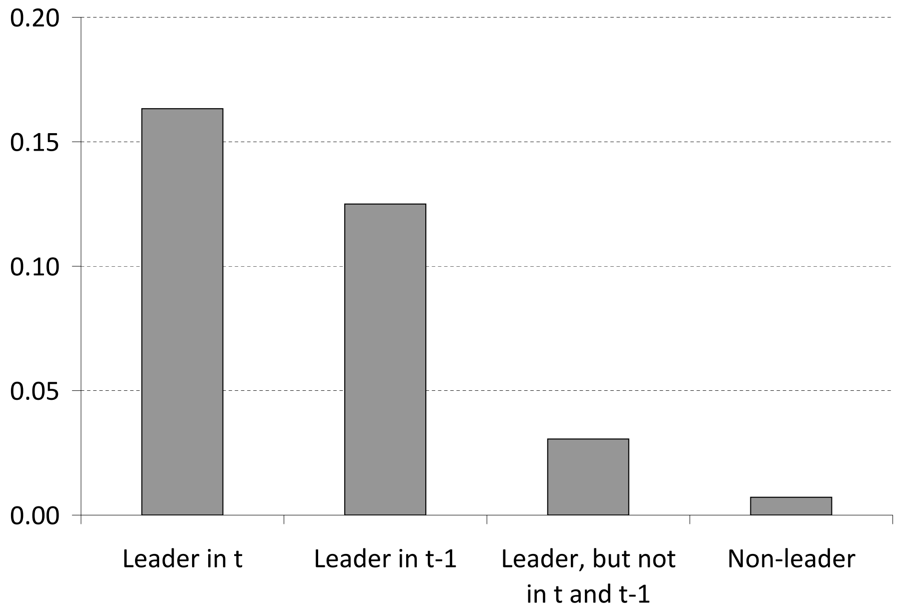

Figure 5: Frequency of payment for giving up anonymity.

Paying for publishing the seat number might not only be due to the leader's desire to become publicly known as a leader but also be used to strengthen the signaling effect of the leading number. It certainly emphasizes the leading number on the other participants' computer screens if the additional seat number entry is displayed as well. In fact, more participants increase their number after a leading bid with (66 percent) than without (48 percent) publication of the leader's seat number, though not significantly.[7]

Result 4 *Leaders are more likely to reveal their identity than non-leaders. They do so mainly in the periods in which they lead.*

[7]In only 3 of 12 groups we observe both a leading bid with and without the leader using the publicity option.

4.4 Leaders have to think

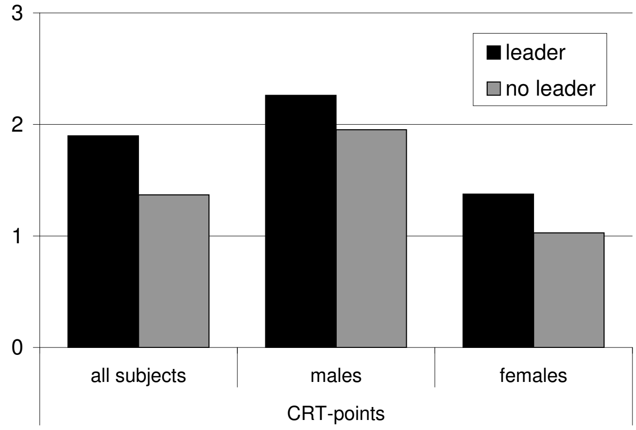

Figure 6: Scores in the cognitive reflection test.

After the main experiment, all participants had to answer three questions from a cognitive reflection test (CRT). There was no incentive for giving a correct answer and no feedback. Each correct answer gives one point in our evaluation so that participants could get between zero and three points in this task. Figure 6 shows that leaders have a significantly higher score on the cognitive reflection test than non-leaders (p-value leaders vs. non-leaders = 0.0773, see also Table 3). The main difference is between leaders and non-leaders, with rather no distinction between early and late leaders.

In line with Frederick (2005), men score higher in the cognitive reflection test than women. In fact, men are significantly more often leaders than women (for a detailed comparison, see again Table 3). Out of the early leaders, 75 percent are males, and 48 percent of the late leaders, while only 37 percent of the non-leaders are men. However the relative difference of the CRT scores for leaders and non-leaders does not depend on gender as can be seen in the second and third pairs of bars in Figure 6 (p-values leaders vs. non-leaders: males = 0.3976, females = 0.0176). Thus, the special kind of intelligence measured by the cognitive reflection test and more frequently observed with men, seems to be a driving force for leading-by-example in the experiment.[8]

Result 5 *Leaders are more willing to use their cognitive capabilities.*

[8]A regression analysis reveals that this result is statistically significant. While male significantly affects the leading probability without further control, it does not when we control for the performance in the CRT task.

Result 6 *Men are more likely leaders. However, this result is driven by the difference between women and men in the cognitive reflection task.*

A second piece of evidence from the recording of reaction times supports this interpretation. Figure 7 illustrates the average time which leaders spend in the belief formation stage and in the decision stage of the experiment. We observe that leaders's belief formation times slow down significantly before their leading decision while the actual decision making gets even faster in the leading round. The regression in Table 4 shows that reaction times of leaders in fact significantly slow down before the decision to lead. "Ever leader?" distinguishes leaders from non-leaders, because it might be that leaders are generally slower or faster than non-leaders in their decisions. The variables "Leader", "Leader in $t+1$?", and "Leader in $t+2$?" are dummy variables being equal to one if the subject is a leader in the respective round. Using them, the regression captures changes in the response times in the leading round and two rounds before compared to rounds in which the person in consideration acts as a leader neither in the current nor in the two subsequent rounds. The formation of beliefs lasts significantly longer in the leading round and already one round before. The time for the actual number choice decision slows down in the two rounds before leading and quickens significantly in the actual leading round.

Our understanding of this change in reaction times is that the decision to lead develops while participants are forced to think about others' behavior in the belief formation stage. The fast leading decision itself seems to be a self-commitment not to rethink the courageous decision to take the initiative.

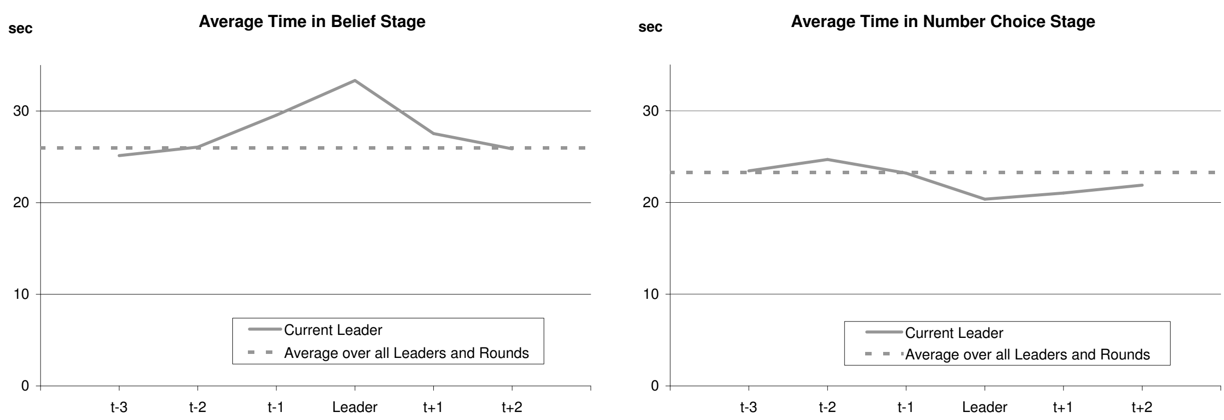

Figure 7: Reaction times.

	log belief time	log decision time
Round	-0.0308***	-0.0114***
	(0.00308)	(0.00169)
Ever Leader?	-0.0211	0.110*
	(0.0620)	(0.0566)
Leader in t?	0.269**	-0.187**
	(0.0989)	(0.0822)
Leader in $t+1$?	0.310***	0.0653
	(0.0922)	(0.0593)
Leader in $t+2$?	0.115	0.0688
	(0.119)	(0.0558)
Constant	3.443***	3.066***
	(0.0710)	(0.0444)

Table 4: Regression coefficients: log of time spend in belief stage and decision stage. Standard errors in brackets. *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level. Standard errors are clustered by matching group.

The reaction times of followers after a leading number further support the idea that reaction times provide a measure for the intensity of thought before a decision. The decision times for the number choice of followers significantly (p-value = 0.0060) slow down in the round after a leading number (23.41 seconds) compared to rounds where no leading number was set in the two previous rounds (20.47 seconds). In the second round after a leading number, average decision times are with 21.83 seconds still slower (p-value = 0.0597). Belief formation times get slower as well. In the round immediately after a leading number, the average belief formation time is 27.60 seconds, compared to 24.33 seconds in rounds without leading number in the two previous rounds (p-value = 0.0597). Two periods after a leading number, the belief formation lasts 28.59 seconds, which is still significantly (p-value = 0.0186) more than in normal rounds.

4.5 Personality measures

None of the personality measures from our questionnaire data has a significant effect on leadership. Out of the big five measures, openness has the strongest correlation coefficient (0.1412). We also cannot find any correlation between an internal locus of control and leadership in our experiment. It seem that leadership is a personality trait that is independent of the established personality traits more related to social preferences.

5 Conclusion

What are the motives behind leadership? We address this question for a particular type of leadership, leading-by-example. We combine an experimental design in which leadership develops endogenously with several other small games and questionnaires to explore possible characteristics and motives of leaders. Our design permits classification of subjects as leaders and non-leaders and to study the determinants of leadership. We find that traditional personality measures are not predictive for leading-by-example. Our main determinants are that leaders attach a high value to efficiency, are not envious, they have better cognitive abilities than non-leaders and they like attention.

Leadership, and in particular leading-by-example is an important and desired trait for many jobs. Thus, measurement devices and the pattern of determinants of this behavior are highly desirable. Our results suggest that traditional personality traits are not very predictive for this behavior. This implies that leadership has to be assessed in a different way. We do not claim that our experiment provides the best or only way to do so. For example, it is not deception proof, and the measure of leadership depends on the comparison group. Nevertheless, it provides interesting insights into the mechanism of leading-by-example and suggests a new way to measure a disposition for leading-by-example, and, in particular, it allows identifying determinants of leadership that are easier to assess.

Appendix: instructions

General instructions at the beginning

Welcome and thank you for participating in this economic experiment.

This experiment consists of multiple parts. The instructions for the first two parts of the experiment will be displayed on your computer screen. The instructions for the third part will be handed out later in hard copy. All instructions are identical for all participants.

Please read the instructions carefully. If you have any questions regarding the experiment please raise your hand. We will then come directly to your place. Please be quiet during the experiment and do not talk to other participants. Failure to comply with these rules will result in an exclusion from the experiment. If this occurs you will not receive any payment.

After you have completed all three parts of the experiment please fill out the following questionnaires on your computer screen. Afterwards you will receive your payment for the entire experiment. The order in which participants receive their payments is already determined, beginning with the participant sitting at the computer "lakelab 1". So take your time to fill in the questionnaires. Your speed will have no influence on the timing of your payment.

Instructions for the number choice game

Now we start with the third part of the experiment. After this part the experiment will be over and we will ask you to fill out some questionnaires.

Your gains and losses during the experiment are counted in points. The exchange rate is 30 points for 1 euro. Your payment in this part of the experiment depends on your decisions and on the decisions of other participants.

This experiment will last for 30 rounds. In each round you will be asked to choose a number between 2 and 100. Subsequently, the computer will randomly determine one participant out of a group of eight and compare the numbers you and the other participant have chosen. The participant who selected the smaller number receives as many points as her number. The other participant receives zero points in this round. If both of you selected the same number, each of you gets half of the points. At the end of each round you are informed about your payment in points and about the numbers all participants of your group have chosen. The composition of your group of eight does not vary during the 30 rounds. Out of this group in each round one participant will be randomly chosen and your numbers will be compared.

In each round before choosing a number you will be asked to make an estimate about the numbers which the other seven participants of your group are going to choose in this round. More specifically, you have to submit your belief about what is going to be the highest, the lowest and the average number of the other seven participants. We ask you to forecast the probability of these three numbers (maximum, minimum and average) being within the following intervals: 2-20, 21-40, 41-60, 61-80 and 81-100. For each of the five intervals you have to indicate the percentage value of the three numbers (maximum, minimum, average) being within these intervals. The five percentage values add up to a total of 100%, because the numbers have to be within one of the intervals no matter what. We place a graphical computer program at your disposal so you can enter your beliefs. You will have the opportunity to familiarize yourself with the program before the experiment begins. Here you can see what the program looks like:

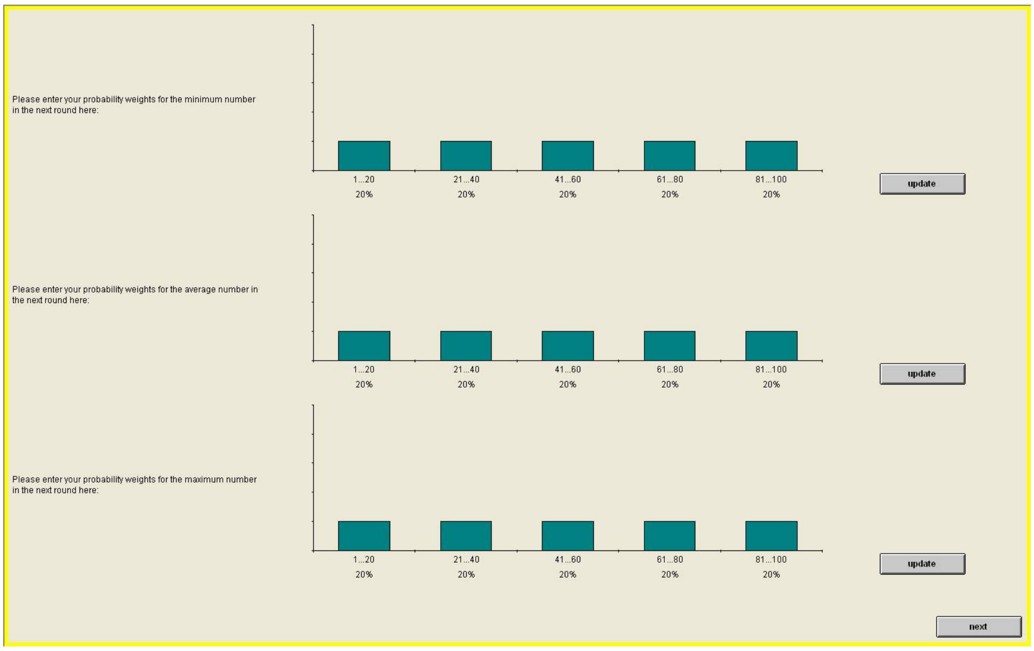

Figure 8: Belief formation tool.

You can change the height of the bars implemented in the program by clicking on a bar, holding the left mouse button and moving the mouse. Do not worry about whether the percentage values add up to 100 or not. Just change the heights of the bars until their proportions match the relative probability you propose. Then click on the button “update” next to the diagram. The bars are automatically adjusted so the values of your estimates sum up to 100. After entering your belief for minimum, maximum and average please click on “next”. Next you can choose your number for the coming round.

There will also be a payment for the accuracy of your guess. The exact computation of this accuracy-dependent payment is described in detail in the appendix. If you have no interest in the details, feel free to ignore the explanations concerning this matter. The only important thing you have to know is that you maximize your payment by indicating your true beliefs.

From the second round on, your previous estimates will be the default setting, so you only have to indicate new numbers in case you want to adjust your previous estimates.

Your decisions in this experiment are always anonymous. The other participants of your group can only see the number you (and all the other participants) have chosen, but not the number of the computer you are sitting at. The numbers are ordered by size. So it is not possible to draw conclusions about participants' seats from the numbers. If in a particular round you want the other participants not only to know the number you have chosen, but also the number of the computer you are sitting at, you can determine so with a mouse click on your computer screen. To disclose the number of your seat you have to pay 10 points.

Before the experiment begins you have the opportunity to familiarize yourself with the computer program. After the experiment please fill in the questionnaires. You will be paid in cash directly after the end of the experiment and after you have finished the questionnaires.

If you have any further questions regarding the conduction of the experiment, please give a short notice to the supervisors of the experiment. We will then come directly to your place.

Payment for probability estimates

As previously described, for the three numbers maximum, minimum and average you allocate five probability values p_i to the five intervals 2-20, 21-40, 41-60, 61-80 and 81-100. The actual number (for example the minimum) lies later in one of these intervals. For one probability estimate you can earn 2 points at most. If your estimate is not accurate there will be subtractions from the 2 points. The probabilities you have assigned to intervals, in which the actual does not lie, will be squared and subtracted from your maximal payment. For example, if you set 70% on the lowest interval but the actual number does not lie in this interval, $0.49 = 0.70 * 0.70$ points will be subtracted from your payment. Furthermore, it is disadvantageous if the probability value you distributed to the interval in which the actual number lies deviates significantly from 100%. This deviation will also be squared and subtracted from your payment. If you set 60% on the right interval, $(1 - 0.60) * (1 - 0.60) = 0.16$ points would be subtracted.

The smaller the sum of the squared wrong estimates is, the better was your guess. For those who are interested, here is the mathematical formula to calculate the quality Q of your guess:

$$Q = 2 - \sum p^2_{wrong,j} - (1 - p_{right})^2$$

In each round the computer will calculate the quality Q of your estimate for minimum, maximum and average number. The higher the quality Q is, the better was your guess in that particular round. At the end of the 30 rounds of the experiment your 30 values of Q for minimum, maximum and average will be summed up. This value will be added to your payment in points.

Examples

In the following we will describe some examples of the calculation of the quality of your estimate and demonstrate some useful tips on how to improve your estimate.

If you think that the smallest of the seven numbers of the other participants of your group definitely is equal to or smaller than 20, you say the probability of the minimum being within the interval 2-20 is 100% and the probability for the minimum being within one of the other intervals is 0%. In this case you gain 2 points if your guess is correct and no points are subtracted for false estimations, because you were 100% right. If you had distributed 20% to each of the five intervals, you would have scored only 1.2 points. In general: if you are sure about the actual number not being within a certain interval, it is better for you to assign a probability of 0% to this interval. Intentional probability "dispersion" does not pay off.

If you think that the highest of the seven numbers of the other participants of your group is either in the interval 61-80 or is higher than 80, but you are sure that the maximum definitely lies above 60, you should assign the value 50% to both intervals 61-80 and 81-100. In this case, your expected payoff is higher than in case you assigned 100% probability to only one of the intervals: If you assign 50% to both of the intervals, you surely gain 1.5 points. If you assigned the entire 100% to one of the intervals, you gained 2 points in case you were right and 0 points in case you were wrong. So your expected payoff would be only 1 point. In general: If you think that the sought number possibly is within several intervals and the probability of the number being in each of these intervals is equal, it is best for you to enter equal probabilities to these intervals.

References

Arbak, E., and M.-C. Villeval. Endogenous Leadership: Selection and Influence. IZA Discussion Paper No. 2732, Institute for the Future of Labor, Bonn, 2007.

Bass, B.M. and Avolio, B.J. (Eds.). (1994). Improving organizational effectiveness through transformational leadership. Thousand Oaks, CA: Sage Publications.

Brandstätter, H. (1988). Sechzehn Persönlichkeits-Adjektivskalen (16PA) als Forschungsinstrument anstelle des 16 PF, Zeitschrift fr experimentelle und angewandte Psychologie 35(3), 370-391.

Brier, G.W. (1950). Verification of Forecasts Expressed in Terms of Probability. Monthly Weather Review 78(1), 1-3.

Bruttel, L. (2009). Group Dynamics in Experimental Studies - The Bertrand Paradox Revisited, Journal of Economic Behavior and Organization 69(1), 51-63.

Dufwenberg, M. and U. Gneezy (2000). Price Competition and Market Concentration: An Experimental Study, International Journal of Industrial Organization 18, 7-22.

Engelmann, D. and M. Strobel (2004). Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments. American Economic Review 94(4), 857-869.

Fischbacher, U., 2007. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Experimental Economics 10, 171-178.

Frederick, S. (2005). Cognitive Reflection and Decision Making, Journal of Economic Perspectives 19(4), 25-42.

Gächter, S., D. Nosenzo, E. Renner, and M. Sefton (forthcoming). Who Makes a Good Leader? Cooperativeness, Optimism and Leading-by-Example, Economic Inquiry.

Gächter, S. and E. Renner (2006), Leaders as Belief Managers for the Voluntary Provision of Public Goods, Working Paper.

Greiner, B. (2004). The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics, University of Cologne, Working Paper Series in Economics 10.

Güth, W., M. V. Levati, M. Sutter, and E. Van Der Heijden (2007). Leading by Example with and without Exclusion Power in Voluntary Contribution Experiments, Journal of Public Economics 91, 1023-1042.

Holt, C.A. and S.K. Laury (2002). Risk Aversion and Incentive Effects, American Economic Review 92, 1644-1655.

Kouzes, J.M., and Posner, B.Z. (2007). The Leadership Challenge, Wiley, San Francisco, CA.

Lerner, J. and J. Tirole (2002). Some Simple Economics of Open Source, Journal of Industrial Economics 50(2), 197234.

Levati, M. V., M. Sutter, and E. van der Heijden (2007). Leading by Example in a Public Goods Experiment with Heterogeneity and Incomplete Information, Journal of Conflict Resolution 51, 793-818.

Moxnes, E. and E. van der Heijden (2007). The Effect of Leadership in a Public Bad Experiment, Journal of Conflict Resolution 51, 793-818

Pogrebna, G., D.H. Krantz, C. Schade, and C. Keser (2009). Leadership by Example and by Pre-Game Communication in Social Dilemma Situations, Working Paper.

Potters, J., M. Sefton, and L. Vesterlund (2007). Leading-by-Example and Signaling in Voluntary Contribution Games: An Experimental Study, Economic Theory 33, 169-182.

Ross, L., D. Greene, and P. House (1977). The "False Consensus Effect": An Egocentric Bias in Social Perception and Attribution Processes. Journal of Experimental Social Psychology 13, 279-301.

Rubinstein A. (2007). Instinctive and Cognitive Reasoning: A Study of Response Times. The Economic Journal 117(523), 1243-1259

Yukl, G. (2009). Leadership in Organizations, Prentice Hall, Upper Saddle River, New Jersey, USA

Zizzo, D.J. (2010). Experimenter Demand Effects in Economic Experiments, Experimental Economics 13(1), 75-98.