No. 125

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

Gerald Eisenkopf, Torben Kölpin

Leading-by-Example: A meta-analysis

Research Paper Series
Thurgau Institute of Economics
and Department of Economics
at the University of Konstanz

Leading-by-Example: A meta-analysis

Gerald Eisenkopf * and Torben Kölpin

Abstract

We provide a parsimonious model of leadership in social dilemma situations and test it with a meta-analysis of experimental studies. We focus on studies with treatments that allow for sequential contributions to a public good (as in Güth et al. (2007)). The group members observe the contribution of a leader before contributing themselves. We compare the results with simultaneous contribution treatments from the same studies. Our results confirm that the establishment of a leader indeed leads to persistently higher and more coordinated contributions. As predicted, the aggregate effect remains stable over time and increases in group size even though leaders and followers have more divergent contribution patterns in larger groups. We also find empirical support for an explanation of the observed 'leader's curse'.

Keywords: *Leading-by-Example, Cooperation, Meta-analysis, Voluntary contribution*

JEL Classification Numbers: *C92, C71, B40*

* Corresponding author: Gerald Eisenkopf, University of Vechta, 49377 Vechta (Germany). Phone: +49 4441 15 127. Fax: +49 4441 15 67211. E-Mail: Gerald.eisenkopf@uni-vechta.de

I. INTRODUCTION

The dynamic interaction between leaders and followers in groups complicates the evaluation of any leadership process. It is difficult to identify causal relationships between the characteristics of leaders and followers, their decisions and group performance. Even with large sample sizes, statistical relationships are often just correlations that do not allow for causal inference because of endogeneity concerns. Hence, leadership research has focused in recent years on experiments as a complimentary research method. A well-designed experimental study eliminates many statistical concerns in field studies (Antonakis et al. 2010) and circumvents related problems of selectivity. Ideally, such experiments take place in natural environments, but such circumstances often reflect idiosyncratic characteristics of the respective context. A clean identification strategy requires carefully controlled and replicable experiments with a design that tests the hypotheses of a specific leadership theory (Falk and Heckman 2009).

We exploit the benefits of this scientific approach in this paper. We focus on experimental studies that use exemplary leadership in a social dilemma (as described in Güth et al. (2007)). More specifically, a randomly selected leader visibly commits herself to a specific contribution to a public good before the fellow group members do so. Our paper studies four questions about the behavioral impact of such leadership. For each question we propose a hypothesis that derives from a parsimonious theory which combines standard economic reasoning with the simple reciprocity model proposed by Fehr and Schmidt (1999). We test this theory with a meta–analysis of experimental studies that allow for clean and replicable identification.[1]

In most studies, the decision sequence of the leadership game increases aggregate contributions and welfare relative to groups without a leader (Moxnes and van der Heijden 2003,

[1] Note that more recent models of reciprocity such as Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006), Cox et al. (2007) seem to provide a better explanation for reciprocal behavior in a variety of games. Our own calculations suggest that these more complex models do not provide qualitatively different results for the leadership game we consider. Hence we apply Occam's razor and stick to the simpler model of Fehr and Schmidt (1999).

Güth et al. 2007, Dannenberg 2015, McCannon 2018, Eichenseer 2021) but this positive effect cannot always be confirmed (Haigner and Wakolbinger 2010, Sahin et al. 2015, Gächter and Renner 2018, Gürerk et al. 2018). Therefore, our first research question focuses on the aggregate impact of such leadership on contributions across several studies.

Question 1: *Does Leading-by-Example increase contributions?*

Aggregate contributions are the most popular indicator of leadership effectiveness, but they do not constitute the only one. Our second research question focuses on the impact of such leadership on coordination. Fischbacher et al. (2001) show that many people are conditional cooperators, who prefer to match others' contributions. If followers are conditional cooperators, leaders' high efforts can influence the followers to exert high efforts. Gächter et al. (2012), Frackenpohl et al. (2016) and Cartwright and Patel (2010) elicited followers contributions using the strategy method (Selten 1967). Their results show a significant correlation between the contributions of followers and leaders. Hence, a cautious leader also has an impact. Their low contributions can decrease the expenditure of followers. Note that a limited degree of reciprocity among followers implies costs and risks for the leader. Some studies report that leaders end up worse than followers do (Cappelen et al. 2016, Gächter and Renner 2018, Eisenkopf 2020) and they may receive even lower payoffs than in a group without leadership. Here, the random allocation into experimental roles is particularly helpful for evaluation because voluntary leadership is a highly selective process that attracts only persons with specific characteristics (Arbak and Villeval 2013, Alan et al. 2019).

Question 2: *Does Leading-by-Example induces a stronger alignment of group members' contributions?*

Third, we focus on the stability of any leadership effect over time. If leaders are also conditional cooperators they might be unwilling to uphold high levels of contribution in case of lower. Gächter and Renner (2018) as well as Teyssier (2012) show that leaders behave almost perfectly like conditional cooperators and match their contribution with the amount they believe the followers will contribute. Consequently, the positive effect of the Leading-by-Example would have to disappear if followers permanently undercut the contribution of the leader.

Question 3: *Do higher contributions with Leading-by-Example persist over time?*

Last, not least, we have an interest in the role of group size on the impact of leadership. Several public good experiments show a positive correlation between group size and contributions in related experiments without a leader (Weimann et al. 2019, Goeree et al. 2002, Isaac and Walker 1988, Isaac et al. 1994). Even if this effect is not robust for all comparisons (Nosenzo et al. 2015, Carpenter 2007), a meta study with 27 experiments confirm the positive correlation (Zelmer 2003). However, only few studies deliberately investigate the role of group size in the context of leadership. Some studies suggest that the effect of Leading-by-Example is also present in larger groups (Figuieres et al. 2012), but the coordination effect of leaders seems to diminish with group size (Komai and Grossman 2009). Therefore, our last research question focuses on the impact of group size on Leading-by-Example.

Question 4*: Does the impact of Leading-by-Example change with the size of the group?*

We obtained data from 14 studies with 369 groups as independent observations. Our results show that Leading-by-Example significantly increases contributions in comparison with leaderless settings. Followers reply in kind to the leader, but only to a certain extent. Consequently, leaders contribute significantly more than followers. Therefore, we can confirm the 'leader's curse' (Gächter and Renner 2018): They earn less than their fellow group members. Nevertheless, leaders do not reduce their contributions more than those of their followers. The positive effect of Leading-by-Example is maintained over a longer period of time. Moreover, we find that contributions increase in group sizes. However, leaders in larger groups elicit less coherent responses from their followers.

Meta-analyses are popular tools in leadership studies from outside economics. Jong et al. (2016) observe in their study (which combines results from 112 studies) that trust between leaders and followers correlates significantly with team performance. Simons et al. (2015) find that the behavioral integrity of a leader has rather high correlations with trust, in-role task performance and organizational citizenship behavior. Their analysis relied on 35 samples. Similarly, Bedi et al. (2016) as well as Zhang et al. (2019) provide meta-analyses on the relationship between ethical leadership and follower outcomes. However, such aggregation of evidence still provides just correlational evidence and does eliminate the methodological limitations of field studies that do not exploit random variations in their identification strategies (Antonakis et al. 2010). Economists rarely use meta-analyses but the increasing popularity of experimental methods with standardized games has provided some studies on public goods games (Croson and Marks 2000, Zelmer 2003), trust games (Johnson and Mislin 2011), ultimatum games (Larney et al. 2019), dictator games (Engel 2011) and the experimental paradigm of Fischbacher and Föllmi-Heusi (2013) on truth-telling (Abeler et al. 2019). More often, economists review the relevant journals in dedicated outlets like the Journal of Economic Literature or the Journal of Economic Surveys. Regarding the topic of our paper, Eichenseer (2021) provides a thorough and thoughtful discussion of the relevant literature and interesting variations of the leadership game. That paper also includes an aggregate quantitative synthesis of the reported results, while our paper allows for an in-depth analysis of the interaction between individual leaders and followers within the groups.

The remainder of the paper is organized as follows. Section II explains the underlying game structure, our leadership model as well as the resulting theoretical predictions. Section III explains the methodology. Section IV presents our main results, while section V concludes.

II. THE UNDERLYING GAME AND THEORETICAL PREDICTIONS

Our study focuses on behavior in the following voluntary contribution game (VCM) that has been by introduced by Isaac and Walker (1988) and adapted to leadership by Güth et al. (2007).

Let $I = (1, \ldots, N\}$ denote a group of $N \geq 3$ individuals who interact for $t = 1, \ldots, T$ periods. In each period t, individual $i \in I$ gets an endowment $e > 0$, which can be either privately consumed or contributed to a group activity. For our theoretical analysis, we set $T = 1$, standardize the endowment $e = 1$ and consider a binary decision regarding the contribution: $c_i \in \{0,1\}$. This simplification allows us to focus on the conditions that induce group members to make nonnegative contributions and how Leading-by-Example alters these conditions. The monetary payoff of individual i takes the following form:

$$\pi_i = 1 - c_i + q \sum_{j=1}^{N} c_j$$

We have an interest in all studies in which $1 > q > \frac{1}{N}$ holds. Because of $1 > q$ the dominant strategy for each rational and selfish player is to contribute nothing. However, because of $q > \frac{1}{N}$, full contributions would generate the highest aggregate payoff of all group members.

We consider two variants of this game: the standard simultaneous VCM and the VCM with leadership. In the simultaneous VCM, all N group members make their contribution decisions privately and simultaneously. The VCM with leadership has two decision stages. First, the leader, L, chooses his contribution c_L, which is observed by the followers. Then, the followers F $(\neq L)$ decide simultaneously about their own c_F. Applying backward induction and assuming commonly known monetary payoff maximization, the theoretical prediction for the VCM with leadership do not differ from those for the standard-VCM: Because of $q < 1$, the followers' dominant strategy in stage 2 is to contribute zero. A rational leader will anticipate this and free-ride as well in stage 1.

However, we want to inquire whether reciprocal preferences induce more cooperation in the group and how Leading-by-Example fosters this cooperation. We assume that players suffer psychological losses from both advantageous and disadvantageous inequality (Fehr and Schmidt 1999).[2] The parameter β_i (with $1 > \beta_i \geq 0$) measures the utility loss from advantageous inequality, while α_i with $\alpha_i \geq \beta_i$ indicates the loss from disadvantageous inequality. For simplicity, we assume that people have common knowledge about a homogeneous $\alpha_i = \alpha > (1-q)$ in the population. Moreover, regarding advantageous inequality aversion we consider only two types of persons ($\beta_i \in \{\underline{\beta}, \overline{\beta}\}$, with $\underline{\beta} = 0$ and $1 - q < \overline{\beta} \leq \alpha$). While the share of $\underline{\beta}$ in the entire population is common knowledge, individual realizations of β_i constitute private knowledge. Groups consist of randomly drawn samples from that population. Let x denote the expected share of people with $\beta_i = \underline{\beta}$ within a group (and $1 - x$ for $\beta_i = \overline{\beta}$). In case of simultaneous contributions, the utility function of a group member can be denoted as follows:

$$U_i = 1 - c_i + q_i + q_j \sum_{j=1}^{N-1} c_j - \frac{\alpha}{N-1}\left(\sum_{j=1 \neq i}^{N-1} max\{c_i - c_j, 0\}\right)$$
$$- \frac{\beta_i}{N-1}\left(\sum_{j=1 \neq i}^{N-1} max\{c_j - c_i, 0\}\right)$$

FS have identified in proposition 4 (with some differences in notation) two conditions for potential asymmetric equilibria in which some group members increase their utility with a contribution. First, for any such member $\beta_i > (1-q)$ must hold. Hence, only people with $\beta_i = \overline{\beta}$ will contribute. Second, the benefits of a contribution must outweigh the costs of free-riding:

[2] Note also that FS provide a more detailed analysis regarding more differentiated contribution possibilities and social preferences.

$$q + q(1-x)(N-1) - \alpha x > 1 + q(1-x)(N-1) - \overline{\beta}\,(1-x)$$

This inequality implies the following condition:

$$\hat{x}_{Sim} = \frac{q - 1 + \overline{\beta}}{\alpha + \overline{\beta}} > x$$

The term $\hat{x}_{Sim}$ denotes the maximum share of expected free riders at which contributions become profitable for conditional cooperators. We now study the case of the sequential VCM. Applying backward induction, we first look at the utility function of a follower F_i who has observed the decision of leader L:

$$U_{F_i} = 1 - c_{F_i} + qc_{F_i} + qc_L + q\sum_{j=1\neq i}^{N-2} c_j$$
$$- \frac{\alpha}{N-1}\left(\max\{c_{F_i} - c_L, 0\} + \sum_{j=1\neq i}^{N-2} max\{(c_{Fi} - c_{Fj}, 0\}\right)$$
$$- \frac{\beta_{Fi}}{N-1}\left(\max\{c_L - c_{F_i}, 0\} + \sum_{j=1\neq i}^{N-2} max\{c_{Fj} - c_{F_i}, 0\}\right)$$

In case of $c_L = 0$, the leader has already revealed her free riding. Again, the benefits of a follower's contribution must outweigh the costs of free-riding for any follower with $\beta_i = \overline{\beta}$:

$$q + q(1-x)(N-2) - \alpha\left(\frac{1 + x(N-2)}{N-1}\right) > 1 + q(1-x)(N-2) - \overline{\beta}\frac{(1-x)(N-2)}{N-1}$$

which leads to the following maximum share of free riders in the group for nonnegative contributions:

$$(\hat{x}_{Seq}|c_L = 0) = \frac{\left(q - 1 - \frac{\alpha}{N-1} + \frac{\overline{\beta}(N-2)}{N-1}\right)}{(\alpha + \overline{\beta})}\left(\frac{N-1}{N-2}\right) < \frac{q + \overline{\beta} - 1}{\alpha + \overline{\beta}} = \hat{x}_{Sim}$$

If the leader has contributed her endowment instead ($c_L = 1$), the number of potential free riders within the group has decreased by one person. Hence, a conditionally contributive follower tolerates a higher share of free riders among the other followers than in the case of simultaneous contributions:

$$\left(\hat{x}_{Seq} \middle| c_L = 1\right) = \left(\frac{q + \overline{\beta} - 1}{\alpha + \overline{\beta}}\right)\left(\frac{N-1}{N-2}\right) > \frac{q + \overline{\beta} - 1}{\alpha + \overline{\beta}} = \hat{x}_{Sim}$$

Now we study the benefit of a contribution for the leader. She will contribute if

$$(U_L | c_L = 1) > (U_L | c_L = 0)$$

holds which depends on the response of the followers towards the choice of the leader. As shown above, this response depends on the realization of x. We have three cases to distinguish:

- $x < \left(\hat{x}_{Seq} \middle| c_L = 0\right)$. In this case, there are enough conditional cooperators in the population. These people will always contribute irrespective of the choice of the leader. Any leader with $\beta_L = \overline{\beta}$ will contribute.
- $\left(\hat{x}_{Seq} \middle| c_L = 1\right) < x$. The followers will never cooperate irrespective of the choice of the leader.
- $\left(\hat{x}_{Seq} \middle| c_L = 0\right) \leq x < \left(\hat{x}_{Seq} \middle| c_L = 1\right)$. In this case, conditional cooperators follow the choice of the leader. This case implies $(U_L | c_L = 1) = q + q(N-1)(1-x) - x\alpha$ while $(U_L | c_L = 0) = 1$ holds. Hence, we obtain

$$x < \frac{qN - 1}{\alpha + q(N-1)} = \hat{x}_{Seq_L}$$

This inequality has an important implication because it provides an explanation why leaders accept lower average payoffs than followers (the leader's curse). The contribution decision in this critical case does not depend on β_L, the 'compassionate' part of a leader's inequality aversion. It rests on the expectation about how many followers the leader can induce to reciprocate, such that the expected net returns from the aggregated investments compensate the expected loss from disadvantageous inequality. This insight implies that leaders should not stop contributions if it pays off even if some group members are free-riders. Comparing $\hat{x}_{Seq_L}$ with $\hat{x}_{Sim} = \left(\frac{q+\bar{\beta}-1}{\alpha+\bar{\beta}}\right)$, the threshold in case of simultaneous contributions, it becomes obvious that a person is more likely to contribute as a leader if $\bar{\beta}$ does not exceed q or otherwise, if the group size is sufficiently large $\left(N > \frac{\bar{\beta}}{q} + 1\right)$.

Hence, our analysis has identified two sources of increased contributions via leadership. Conditional cooperators are more likely to contribute if the leader has contributed herself while assuming the leadership role itself tends to induce contributions in particular among otherwise uncooperative group members. Overall, our theoretical analysis suggests the following expectations regarding the research questions:

Hypothesis 1*: Leading-by-Example increase contributions.*

Hypothesis 2: *Leading-by-Example induces a stronger alignment of group members' contributions.*

Hypothesis 3: *Higher contributions with Leading-by-Example persist over time.*

Group size has an ambiguous effect in this context. If x is in the appropriate parameter range, an increase in group size makes a leader's contribution more rewarding. At the same time, an increasing group size decreases that parameter range, as the leader's decision becomes less relevant for the fellow group members. However, this result relied on the assumption about common knowledge regarding the share of people with $\beta_i = \underline{\beta}$. This assumption is rather bold in the context of anonymous interaction between participants who do not know each other. Hence, followers will have a prior about the share. They can use a leader's decision to update this prior. Therefore, a contribution by the leader can increase the expectations of the followers regarding the share of participants with $\bar{\beta}$. This revision of expectations does not depend on the group size.

Hypothesis 4: *For a given level of marginal per capita return (the variable q), larger groups will see higher investments of leaders than smaller groups.*

III. METHOD

III.A Search of Studies and Study criteria

We searched for relevant studies in early 2020. More specifically, we first looked for experimental studies that investigate leadership with a voluntary contribution mechanism. Leadership in this context means that one of the group members acts first and the others observe the behavior before they act themselves. We applied a three-step search procedure. First, we searched relevant databases for published studies, including e.g., Google Scholar, EconLit or IDEAS, using terms like *Leadership*, *Leading-by-Example, Voluntary Contribution Mechanism, Sequential Contribution, Public Goods*. This first search yielded 251 potential results. We removed all theoretical papers and those that did not use a voluntary contribution mechanism. Second, we checked the references in the remaining papers and looked at their citations in Google Scholar. These steps resulted in potential 33 studies. Last, not least, we sent a request to the e-mail list of the *Economic Science Association (ESA),* asking for additional publications as well as working papers and other unpublished research. The request identified 15 additional studies (including our own relevant working papers). Thus, our search procedure yielded 48 studies.

Table 1: Included Studies

Study	# of Periods	Group Size	MPCR	Endow-ment	Simultaneous Treatment included	Location
Centorrino & Concina (2013)	10	4	0.5	30	---	University of Venice, Italia
Dannenberg (2015)	10	4	0.4	25	✓	University of Magdeburg, Germany
Drouvelis & Nosenzo (2013)	10	3	0.5	20	--	University of Nottingham, England
Eisenkopf & Kölpin (2021)	20	3/6	0.5	100	✓	University of Hamburg, Germany
Eisenkopf (2020)	20	3	0.5	100	✓	University of Konstanz, Germany
Eisenkopf & Walter (2021)	20	3	0.5	100	✓	University of Hamburg, Germany
Frackenpohl et al. (2016)	10	4	0.4	20	--	University of Bonn, Germany
Gächter & Renner (2018)	10	4	0.4	20	✓	University of Erfurt, Germany
Gürerk et al. (2018)	20	4	0.4	20	✓	Aachen University, Germany
Güth et al. (2007)	16	4	0.4	25	✓	Max Planck Institute of Economics, Germany
Moxnes & van der Heijden (2003)	10	3	0.4	20	✓	Norwegian School of Economics, Norway
Rivas & Sutter (2011)	16	4	0.4	25	✓	Max Planck Institute of Economics, Germany
Sahin et al. (2015)	20	6	0.2	9	✓	Virginia Tech & University of Texas, United States
Yu & Kocher (2020)	10	4	0.8/0.4	20	✓	University of Munich, Germany

Note: Eisenkopf & Kölpin (2021) report two Leading-by-Example treatments with different group sizes (3 and 6). Moxnes and van der Heijden (2003) used a public bad environment. We consider only the contribution into the non-damaging good. Yu and Kocher (2020) implement a public goods experiment with heterogeneous marginal per capita returns.

Subsequently, we further narrowed down our criteria. More specifically we looked for studies that met the following criteria: First, to ensure comparable procedures, we restricted the studies to experimental studies that used the voluntary contribution mechanism as in Güth et al. (2007). Second, each relevant study had to have at least one treatment with a randomly allocated leader in order to avoid self-selection among leaders. Third, all participants remained in their groups throughout the entire experiment. Applying these criteria resulted in a total of 20 potential studies for which we requested the data. We obtained the results in 15 cases. However, for one study we only got data on the aggregated level, so we were unable to recover the individual decisions. In total, we received full data from 14 studies, which we include in our meta-study. Table 1 informs about the included studies. We hope that future meta-analyses will be able to replicate our analyses with more samples from more diverse populations.

III.B Coding and Data preparation

For each study, we first transcribed the individual observations into a general form. This data form contained all variables necessary to answer our research questions. Our prime dependent variable was the individual contribution towards to public good. However, since all studies has different maximum stakes, we calculated the individual contribution as a percentage of the respective endowment. This procedure facilitated the comparison of the results of the different studies. Moreover, we included the respective round, the role in the experiment (leader or follower), the subject number as well as the receptive group. Note, that we later assigned a unique identification number to each study, group and subject, to ensure identification. Besides these subject dependent variables, we added study specific characteristics that have been hypothesized in the literature to affect contributions towards a public goods. Such characteristics include the marginal per capita return (mpcr), group size, number of periods, endowment and the exchange rate of the tokens/points into monetary amounts (Ledyard 1995, Zelmer 2003). To ensure comparability, we have converted the exchange rate into euros and adjusted it for inflation (as of summer 2021). Last but not least, we added treatment variables that indicate whether a group played in a sequential or simultaneous contribution structure. After applying the described procedure to all studies, all data sets were transferred into one complete data set.

IV. RESULTS

Before we focus on our research questions in detail, we have a brief look at the descriptive statistics. Table 2 presents the average relative contributions for the simultaneous and Leading-by-Example structures, as well as the number of included studies, groups and subjects. Appendix 0 report additional descriptive statistics for each included study separately. For the sake of clarity, we subdivide the results part in different subsection. Each subsection provides results for at least one of our hypotheses presented in section II. Note that we report results from additional robustness checks in the appendix. More specifically, we looked whether individual studies had an outsized effect on the aggregate outcomes in any of the regression models. We replicated the estimations and eliminated each individual study in a specific subsample for each robustness check. In a few cases, elimination leads to minor changes. The robustness checks are shown in the appendix A.1. We refer to the checks in the corresponding sections.

Table 2: Descriptive Statistics

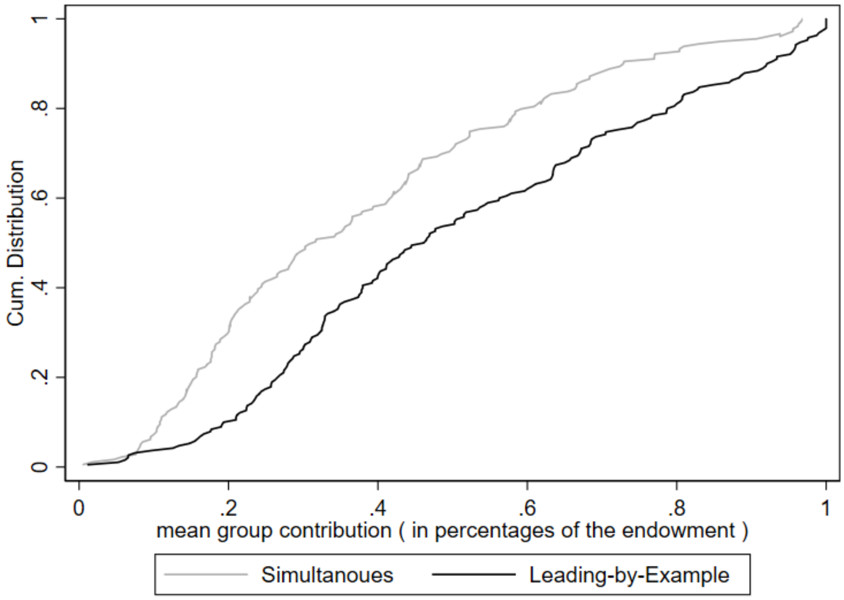
		Simultaneous	Leading-by-Example
Average rel. contributions			
	All	0.379 (0.245)	0.507 (0.2647)
	Leader	---	0.596 (0.276)
	Follower	---	0.478 (0.276)
Studies		11	14
Groups		179	248
Subjects		686	970

Note: Standard Deviations in parentheses

IV.A Does Leading-by-Example increase aggregate cooperation?

To address our first research question, we include only those studies on our subsequent analysis that allow a comparison between sequential and simultaneous contributions (i.e., with and without leader). This leaves us with 11 studies[3], including 369 groups (179 for simultaneous, 190 for Leading-by-Example) as independent observations. Figure 1 demonstrates the cumulative distribution of means from all included groups separated by the contribution (in percentages of the endowment). The distribution for groups with a simultaneous contribution structure is always above the distribution for groups with a leader. A two-sample Kolmogorov-Smirnov test confirms that the distributions are statistically different ($p < .01$).

Figure 1: Cumulative distribution of the mean rel. contribution per group



To address our first research question in greater detail, Table 3 shows the overall impact of Leading-by-Example on contributions. The dependent variable is a subject's contribution to the public good as a percent of the endowment, with standard errors clustered at the group level. The simultaneous decision structure serves as the benchmark. The variable Leading-by-Example denotes the dummy variable for the groups with leaders. Model I studies the impact of Leading-by-Example. The variable is highly significant and indicates a positive impact of Leading-by-Example. The effect of Leading-by-Example remains positive and highly significant

[3] Included studies: Dannenberg (2015), Eisenkopf (2020), Eisenkopf and Kölpin (2021), Eisenkopf and Walter (2021), Gächter and Renner (2018), Gürerk et al. (2018), Güth et al. (2007), Moxnes and van der Heijden (2003), Rivas and Sutter (2011), Sahin et al. (2015), Yu and Kocher (2020).

even if we control for characteristics of the experimental public good environment (model II) or include fixed effects for the studies (model III).

Result 1: *Leading-by-Example enhances cooperation in comparison with simultaneous decisions.*

Table 3: Leading-by-Example in comparison with simultaneous decisions

Dep. Var.: Indiv. contribution in percent of the endowment	**Both Treatments Benchmark: Simultaneous**		
	I	**II**	**III**
Leading-by-Example	0.127***	0.110***	0.110***
	(0.030)	(0.025)	(0.025)
Group size		0.117***	0.092***
		(0.015)	(0.021)
Exchange rate (in €)		-1.491***	-39.413
		(0.360)	(92.601)
MPCR		0.127	0.685***
		(0.124)	(0.103)
Endowment		-0.001**	-0.050
		(0.001)	(0.112)
Total number of periods		0.018***	0.366
		(0.004)	(0.773)
Fixed Effects for Studies	--	--	✓
Constant	0.414***	-0.303***	-2.298
	(0.021)	(0.114)	(3.621)
Observations	23,544	23,544	23,544
Number of Groups	369	369	369
R-squared	0.026	0.137	0.156

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

IV.B To which extent do group members follow the leader?

Next, we focus on the impact of Leading-by-Example on followers. We use the data from groups with leaders of the 11 studies from the previous subsection, but we also include groups from three other studies (Frackenpohl et al. 2016, Centorrino and Concina 2013, Drouvelis and Nosenzo 2013) that did not allow comparison with simultaneous contribution structures but investigated Leading-by-Example in other contexts. In appendix 0 we provide robustness checks without the three additional studies. Excluding these three studies does not alter the results significantly. Table 4 presents OLS-regressions with the individual contribution as a percent of the endowment as the dependent variable. We include the characteristics of the experimental public good environment as independent variables. In addition, the dummy variable *Leader not fixed* indicates whether the leader remains in her role during the experiment (=0) or whether the leader role is changed between periods (=1). Models I - III address contribution differences between leaders and followers. These models include a dummy variable which indicates whether a subject is in the role of a leader (=1) or a follower (=0). All three models show a positive and highly significant coefficient for the *Leader* variable. Thus, leaders contribute significantly more to the public good than followers. Moreover, contributions in groups with a fixed leader are higher than in groups with different leaders.

Table 4: The impact of Leading-by-Example

Dep. Var.: Indiv. contribution in percent of the endowment	**Leading-by-Example**					
	Leader & Follower			**Only Followers**		
	I	**II**	**III**	**IV**	**V**	**VI**
Leader	0.090*** (0.012)	0.112*** (0.010)	0.111*** (0.010)			
Rel. Leader contribution				0.676*** (0.023)	0.692*** (0.023)	0.705*** (0.022)
Group Size		0.076*** (0.019)	0.074** (0.029)		-0.002 (0.013)	-0.007 (0.016)
Exchange rate (in €)		0.471 (0.374)	-3.114 (8.550)		-0.245 (0.275)	5.380 (4.512)
MPCR		0.446** (0.174)	0.653*** (0.127)		0.495*** (0.152)	0.794*** (0.157)
Endowment		-0.000 (0.001)	-0.012* (0.007)		-0.001** (0.001)	-0.002 (0.005)
Total number of periods		0.014*** (0.005)	0.042 (0.068)		0.016*** (0.004)	-0.027 (0.035)
Leader not fixed	-0.161*** (0.038)	-0.134*** (0.044)	-0.166** (0.073)	-0.056** (0.027)	-0.071** (0.028)	-0.099*** (0.033)
Fixed Effects for Studies	---	---	✓	---	---	✓
Constant	0.531*** (0.021)	-0.231* (0.136)	-0.231 (0.425)	0.089*** (0.013)	-0.317*** (0.111)	-0.098 (0.233)
Observations	14,392	14,392	14,392	10,676	10,676	10,676
Number of Groups	248	248	248	248	248	248
R-squared	0.028	0.080	0.108	0.396	0.419	0.461

OLS-Regression. Robust standard errors in parentheses clustered at group level.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

To identify the impact of the example set by the leader, model IV - VI in Table 4 include only observations of the followers. Again, the dependent variable is the individual contribution as a percent of the endowment. The variable *Rel. Leader contribution* indicates the contribution of the leader in a given period. In all estimation models, this variable is positive and highly significant. However, both models indicate that followers, on average, employ an imperfect matching strategy.

The previous findings confirm that Leading-by-Example influences follower behavior. Our second hypothesis predicts that this impact leads to a stronger alignment of the group members' contributions. Thus, the heterogeneity of individual contributions should be smaller in groups with a leader. Table 5 shows the alignment of group members contributions in detail. The estimations reported in Table 5 use the data from the 369 groups of the 11 studies that allow for a comparison between Leading-by-Example and simultaneous contributions. We use the standard deviation within a group in each period as measure of intra-group heterogeneity. The variable *Leading-by-Example* denotes a dummy variable which indicates the groups with leaders. All three models show that the standard deviation within groups is significantly lower in the Leading-by-Example treatments in comparison with simultaneous contributions. Leading-by-Example leads to a stronger alignment of contributions within the group. This observation confirms our second hypothesis. Note, however, that the effect vanishes when we exclude Eisenkopf and Walter (2021) or Güth et al. (2007) (see appendix 0).

Results 2: *Leading-by-Example leads to a stronger alignment of the contributions of the group members.*

Table 5: Standard deviations of individual contributions within groups

Dep. Var.: Standard deviations within groups	**Both Treatments Benchmark: Simultaneous**		
	I	**II**	**III**
Leading-by-Example	-0.020*	-0.029***	-0.029***
	(0.011)	(0.009)	(0.010)
Group Size		0.039***	0.040***
		(0.010)	(0.012)
Exchange rate (in €)		0.170	-30.799
		(0.206)	(32.870)
MPCR		0.122	0.037
		(0.076)	(0.091)
Endowment		-0.000	-0.040
		(0.000)	(0.040)
Total number of periods		-0.004**	0.247
		(0.002)	(0.274)
Fixed Effects for Studies	---	---	✓
Constant	0.190***	0.074	-0.929
	(0.007)	(0.064)	(1.290)
Observations	6,126	6,126	6,126
Number of Groups	369	369	369
R-squared	0.004	0.126	0.133

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

IV.C What is the long-term impact of Leading-by-Example?

To answer our third research question, Table 6 presents results from an OLS-regression, which focus on long-term effects of Leading-by-Example. The first three models estimate the long-term effect for Leading-by-Example in comparison with simultaneous contribution structures. To estimate the long-term effect, we only include the groups of the 11 studies used in section IV.A. The dependent variable is the individual contribution as a percent of the endowment. *Leading-by-Example* denote the dummy for groups with a leader. Model II controls for characteristics of the experimental public good environment, while model III includes fixed effects for studies. Our estimations show that contributions generally decrease over time. To test the long-term differences between groups with and without leaders, we implement an interaction term between Leading-by-Example and the period in our estimations. The interaction term enters positively and significantly in all three models. Thus, while during an experiment the contributions generally decrease, the effect is less pronounced in groups with leaders than in groups with simultaneous contributions. Note, however, that this finding is not robust against all specifications. More precisely, the interaction term for model I - III turns insignificant if we exclude Eisenkopf and Walter (2021) from the analysis. However, the Leading-by-Example as well as the period effect remain highly significant. In appendix 0 in Table A. 6 we provide the results from regressions in which we exclude this study.

Table 6: Long-term impact of leadership

Dep. Var.: Indiv. contribution in percent of the endowment	**Both treatments Benchmark: Simultaneous**			**Leading-by-Example**		
	I	**II**	**III**	**IV**	**V**	**VI**
Leading-by-Example (LbE)	0.108***	0.082***	0.080***			
	(0.0320)	(0.0284)	(0.0294)			
LbE*Period	0.0050**	0.0049**	0.0049**			
	(0.0024)	(0.0023)	(0.0022)			
Leader				0.097***	0.11***	0.104***
				(0.0171)	(0.0162)	(0.0155)
Leader*Period				-0.0008	0.0003	0.0008
				(0.0013)	(0.0012)	(0.0011)
Period	-0.009***	-0.012***	-0.012***	-0.0023	-0.007***	-0.007***
	(0.0017)	(0.0014)	(0.0014)	(0.0018)	(0.0015)	(0.0015)
Group Size		0.116***	0.092***		0.076***	0.074**
		(0.0152)	(0.0216)		(0.0195)	(0.0289)
Exchange rate (in €)		-1.502***	-43.1579		0.4705	18.10***
		(0.3651)	(93.0367)		(0.3739)	(5.9670)
MPCR		0.2587**	0.685***		0.4455**	0.653***
		(0.1246)	(0.1028)		(0.1737)	(0.1268)
Endowment		-0.0016**	-0.0510		-0.0004	0.0006
		(0.0006)	(0.0936)		(0.0009)	(0.0047)
Total number of periods		0.022***	0.3983		0.018***	-0.1290**
		(0.0042)	(0.7974)		(0.0055)	(0.0504)
Leader not fixed	-0.185***	-0.125***	-0.1161**	-0.166***	-0.134***	(omitted)
	(0.0391)	(0.0433)	(0.0588)	(0.0377)	(0.0443)	
Fixed Effects for Studies	---	---	✓	---	---	✓
Constant	0.494***	-0.297***	-2.3872	0.552***	-0.2264*	0.6710
	(0.0214)	(0.1114)	(4.2298)	(0.0220)	(0.1367)	(0.4268)
Observations	23,544	23,544	23,544	14,392	14,392	14,392
Number of Groups	369	369	369	248	248	248
R-squared	0.0482	0.1606	0.1761	0.0292	0.0883	0.1154

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

Model IV-VI now focus on the long-term effect in groups with leaders. For our estimations we rely on the groups of the 14 studies used in section IV.B. Again, section 0 provides robustness checks that exclude the three additional studies. The variable *Leader* denotes whether a participant held the role of leader (=1) or a follower (=0). Model V controls for characteristics of the experimental public good environment, while model VI additionally control for the studies. Again, our models show that contributions decrease during an experiment. Leaders generally contribute more than followers. However, we do not find that leader and followers do react differently to the progress of the experiment which is in line with our third hypothesis.

Result 4: *Leading-by-Example also has a positive effect on contributions in the long-term view. Leaders do not decrease their contributions more than followers.*

IV.D Do contributions increase with group size?

Last not least, we focus on group sizes effects on contributions. Table 7 presents results from an OLS-regression based on the groups of the 11 studies that compare Leading-by-Example and simultaneous contribution structures (see section IV.A). The dependent variable is the individual contribution to the public goods, with standard errors clustered at the group level. Model I shows that contributions increase with larger groups, but that contributions are higher in the Leading-by-Example treatments. Model II examines the differential effects of group size on the two treatments. It suggests a positive effect of group size, but indicates that this effect is smaller in groups with a leader. These findings remain highly significant even when we control for the public goods environment (model III) and include study fixed effects (model IV).

Table 7: The impact of group size

Dep. Var.: Indiv. contribution in percent of the endowment	**Both treatments** **Benchmark: Simultaneous** **I**	**II**	**III**	**IV**
Leading-by-Example (LbE)	0.109*** (0.026)	0.328*** (0.095)	0.346*** (0.093)	0.370*** (0.093)
Group Size	0.093*** (0.012)	0.122*** (0.014)	0.148*** (0.016)	0.126*** (0.023)
Group Size * LbE		-0.054** (0.022)	-0.058*** (0.021)	-0.063*** (0.021)
Exchange rate (in €)			-1.469*** (0.365)	-43.893 (92.826)
MPCR			0.115 (0.125)	0.685*** (0.103)
Endowment			-0.001* (0.001)	-0.055 (0.094)
Total number of periods			0.018*** (0.004)	0.406 (0.796)
Fixed Effects for Studies	---	---	---	✓
Constant	0.042 (0.048)	-0.075 (0.059)	-0.432*** (0.119)	-2.655 (4.223)
Observations	23,544	23,544	23,544	23,544
Number of Groups	369	369	369	369
R-squared	0.097	0.103	0.144	0.164

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

To examine the effects of group size on Leading-by-Example in more detail, we focus below on the groups of the 14 studies that focused on Leading-by-Example (see section IV.B). Again, section 0 provides robustness checks that exclude the three additional studies. Table 8 presents results from an OLS-regression. Model I – III focus on the leaders, while the remaining models consider the followers. The first two models confirm our hypothesis 4. Leaders increase their own relative contributions in larger groups. The effect remains significant even if we control for other characteristics of the experimental public good environment (model II) and when we include the fixed effects for studies (model III). Turning to the followers, model IV shows that followers also increase their relative contributions with increasing group sizes. However, the group size variables turn insignificant once we control for the relative leader contribution, whereas the relative leader contribution variable enters positive and highly significant (model V). Model VI and VII identifies the impact of group size on the leader-follower relationship. Model VI shows a group size effect for the followers. However, the effect vanishes if we control for the leader's contribution (model V). Model VI and VII investigate the coordination impact of the leader. Both models show that the coordination impact of leaders on followers becomes weaker for increasing group sizes. However, model I shows that leaders in smaller groups are more timid. Hence, even though the leader has less coordination power in the larger groups she generates more contributions. Note, however, that the group size effect for followers in model IV turns insignificant when we exclude Sahin et al. (2015). All other models do not change significantly (see appendix 0).

Table 8: The impact of group size in Leading-by-Example treatments

Dep. Var.: Indiv. contribution in percent of the endowment	**Leading-by-Example**						
		Leader			**Follower**		
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**
Group Size	0.10***	0.11***	0.10***	0.06***	-0.011	0.044**	0.07***
	(0.015)	(0.018)	(0.027)	(0.017)	(0.011)	(0.019)	(0.025)
Rel. Leader contr. (RLC)					0.69***	0.96***	1.10***
					(0.024)	(0.096)	(0.092)
Group Size * RLC						-0.07***	-0.10***
						(0.025)	(0.025)
Exchange rate (in €)		1.12***	2.32***				-3.04***
		(0.380)	(0.456)				(0.486)
MPCR		0.559**	0.77***				0.79***
		(0.222)	(0.238)				(0.157)
Endowment		-0.000	-0.006				-0.008**
		(0.001)	(0.005)				(0.003)
Total number of periods		0.010*	0.008				0.04***
		(0.006)	(0.008)				(0.006)
Leader not fixed	-0.14***	-0.14***	-0.087	-0.16***	-0.057**	-0.066**	-0.17***
	(0.046)	(0.050)	(0.057)	(0.038)	(0.027)	(0.027)	(0.035)
Fixed Effects for Studies	---	---	✓	---	---	---	✓
Constant	0.23***	-0.243	-0.264	0.26***	0.16***	-0.079	-0.73***
	(0.069)	(0.160)	(0.301)	(0.073)	(0.039)	(0.066)	(0.191)
Observations	3,692	3,692	3,692	10,700	10,676	10,676	10,676
Number of Groups	248	248	248	248	248	248	248
R-squared	0.080	0.097	0.121	0.050	0.397	0.400	0.468

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

The result confirms our theoretical prediction from hypothesis 4. Rather reluctant leadership in small groups explain this gap even though leaders in large teams elicit less coordinated responses from their fellow group members.

Result 5: *Contributions increase with in group size, but the coordination impact of leaders decreases.*

V. DISCUSSION AND CONCLUSION

Experimental studies become more and more important in leadership studies, mainly for three reasons. First, they eliminate endogeneity concerns in the identification of causal relationships. Second, they can be tailored to test a specific theory. Third, they allow for replications by other researchers. We exploited these benefits and merged data from 14 studies in a meta-analysis to answer four questions about the impact of exemplary leadership in the light of a theory that combined standard economic reasoning with a simple model of reciprocity.

First, we hypothesized that Leading-by-Example increases contributions in a social dilemma. The results support this hypothesis. The establishment of a first-moving leader generates significantly higher contributions in comparison to groups without a leader. Our second focus was on the alignment of decisions between leader and followers. Our model predicted that conditional cooperators follow the leader's decision, while the rest refuse to make any contribution. As a result, leaders with high contributions will end up worse than followers (the 'leader's curse'). We observe that leadership generates a greater alignment of group members' contributions even though some followers contribute much less than their leader. Hence, on average, leaders contribute more than followers. We then inquired whether higher contributions with Leading-by-Example persist over time. Our model predicted that even selfish leaders should not stop contributing with sufficiently few likely free riders in their group because their own economic losses from a breakdown in cooperation are too large. The results support these insights. Despite the relatively small gains of leaders, they do not reduce their contributions more than followers over time which establishes a positive long-term effect of Leading-by-Example. These findings highlight the importance of cooperative leadership for successful groups. Last but not least, our fourth question deals with the impact of group size. While our simple model predicts an ambiguous effect, further considerations of Bayesian Updating suggest that the effect of leading by example is stronger in larger groups. Our results show that contributions increase with group size, independent of the contributions structure. This result is consistent with previous literature examining the effects of group size (Weimann et al. 2019, Zelmer 2003, Goeree et al. 2002). At the same time, however, our results show that the effect of a leader decreases as group size increases. Further analysis shows that this is particularly related to the fact that leaders in larger groups elicit fewer coherent responses from their followers. This result suggests that the benefits of Leading-by-Example do not extent beyond a certain group size.

We hope that future meta-analyses can rely on a larger and more diverse sample. Such studies could also test, and potentially falsify, specific extensions of our rather simple leadership model. Moreover, we did not investigate any leadership instruments such as communication, monitoring or punishment. Furthermore, most leaders emerge endogenously within a group or that they come as outsiders into the group. Nevertheless, we consider our results as encouraging because they derive from a systematic, replicable and theory-guided research agenda that may complement and inspire future research in the lab and the 'real life'.

Gerald Eisenkopf
University of Vechta
Driverstraße 22
49377 Vechta (Germany)
Phone: +49 4441 15 127
Fax: +49 4441 15 67211
E-Mail: Gerald.eisenkopf@uni-vechta.de

Torben Kölpin
University of Vechta
Driverstraße 22
49377 Vechta (Germany)
Phone: +49 4441 15 126
Fax: +49 4441 15 67211
E-Mail: torben.koelpin@uni-vechta.de

References

Abeler, J., D. Nosenzo, C. Raymond. 2019. Preferences for Truth-Telling. *Econometrica* **87**(4) 1115–1153.

Alan, S., S. Ertac, E. Kubilay, G. Loranth. 2019. Understanding Gender Differences in Leadership. *The Economic Journal* **130**(626) 263–289.

Antonakis, J., S. Bendahan, P. Jacquart, R. Lalive. 2010. On Making Causal Claims: A Review and Recommendations. *The Leadership Quarterly* **21**(6) 1086–1120.

Arbak, E., M.-C. Villeval. 2013. Voluntary Leadership: Motivation and Influence. *Social Choice and Welfare* **40**(3) 635–662.

Bedi, A., C. M. Alpaslan, S. Green. 2016. A Meta-Analytic Review of Ethical Leadership Outcomes and Moderators. *Journal of Business Ethics* **139**(3) 517–536.

Cappelen, A. W., B.-A. Reme, E. Ø. Sørensen, B. Tungodden. 2016. Leadership and Incentives. *Management Science* **62**(7) 1944–1953.

Carpenter, J. P. 2007. Punishing Free-Riders: How Group Size Affects Mutual Monitoring and the Provision of Public Goods. *Games and Economic Behavior* **60**(1) 31–51.

Cartwright, E., A. Patel. 2010. Public Goods, Social Norms, and Naïve Beliefs. *Journal of Public Economic Theory* **12**(2) 199–223.

Centorrino, S., L. Concina. 2013. A Competitive Approach to Leadership in Public Good Games. *Working Paper*.

Cox, J. C., D. Friedman, S. Gjerstad. 2007. A Tractable Model of Reciprocity and Fairness. *Games and Economic Behavior* **59**(1) 17–45.

Croson, R. T., M. B. Marks. 2000. Step Returns in Threshold Public Goods: A Meta-and Experimental Analysis. *Experimental Economics* **2**(3) 239–259.

Dannenberg, A. 2015. Leading by Example versus Leading by Words in Voluntary Contribution Experiments. *Social Choice and Welfare* **44**(1) 71–85.

Drouvelis, M., D. Nosenzo. 2013. Group Identity and Leading-by-Example. *Journal of Economic Psychology* **39** 414–425.

Dufwenberg, M., G. Kirchsteiger. 2004. A Theory of Sequential Reciprocity. *Games and Economic Behavior* **47**(2) 268–298.

Eichenseer, M. 2021. *Leading by Example in Public Goods Experiments: What Do We Know?*

Eisenkopf, G. 2020. Words and Deeds – Experimental Evidence on Leading-by-Example. *The Leadership Quarterly* **31**(4) 101383.

Eisenkopf, G., T. Kölpin. 2021. Leadership and Cooperation in Growing Teams. *Working Paper*.

Eisenkopf, G., C. Walter. 2021. Leadership Instruments as a Motor of Cooperation – Experimental Evidence on the Impact of Leading-by-Example, Monitoring and Punishment. *Working Paper*.

Engel, C. 2011. Dictator Games: A Meta Study. *Experimental Economics* **14**(4) 583–610.

Falk, A., U. Fischbacher. 2006. A Theory of Reciprocity. *Games and Economic Behavior* **54**(2) 293–315.

Falk, A., J. J. Heckman. 2009. Lab Experiments are a Major Source of Knowledge in the Social Sciences. *Science* **326**(5952) 535–538.

Fehr, E., K. M. Schmidt. 1999. A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics* **114**(3) 817–868.

Figuieres, C., D. Masclet, M. Willinger. 2012. Vanishing Leadership and Declining Reciprocity in a Sequential Contribution Experiment. *Economic Inquiry* **50**(3) 567–584.

Fischbacher, U., F. Föllmi-Heusi. 2013. Lies in Disguise - An Experimental Study on Cheating. *Journal of the European Economic Association* **11**(3) 525–547.

Fischbacher, U., S. Gächter, E. Fehr. 2001. Are People Conditionally Cooperative? Evidence from a Public Goods Experiment. *Economics Letters* **71**(3) 397–404.

Frackenpohl, G., A. Hillenbrand, S. Kube. 2016. Leadership Effectiveness and Institutional Frames. *Experimental Economics* **19**(4) 842–863.

Gächter, S., D. Nosenzo, E. Renner, M. Sefton. 2012. Who Makes a Good Leader? Cooperativeness, Optimism, and Leading-by-Example. *Economic Inquiry* **50**(4) 953–967.

Gächter, S., E. Renner. 2018. Leaders as Role Models and 'Belief Managers' in Social Dilemmas. *Journal of Economic Behavior & Organization* **154** 321–334.

Goeree, J. K., C. A. Holt, S. K. Laury. 2002. Private Costs and Public Benefits: Unraveling the Effects of Altruism and Noisy Behavior. *Journal of Public Economics* **83**(2) 255–276.

Gürerk, Ö., T. Lauer, M. Scheuermann. 2018. Leadership with Individual Rewards and Punishments. *Journal of Behavioral and Experimental Economics* **74** 57–69.

Güth, W., M. V. Levati, M. Sutter, E. van der Heijden. 2007. Leading by Example with and without Exclusion Power in Voluntary Contribution Experiments. *Journal of Public Economics* **91**(5-6) 1023–1042.

Haigner, S. D., F. Wakolbinger. 2010. To Lead or not to Lead: Endogenous Sequencing in Public Goods Games. *Economics Letters* **108**(1) 93–95.

Isaac, R., J. M. Walker, A. W. Williams. 1994. Group Size and the Voluntary Provision of Public Goods. *Journal of Public Economics* **54**(1) 1–36.

Isaac, R. M., J. M. Walker. 1988. Communication and Free-Riding Behavior: The Voluntary Conbtribution Mechanism. *Economic Inquiry* **26**(4) 585–608.

Johnson, N. D., A. A. Mislin. 2011. Trust games: A Meta-Analysis. *Journal of Economic Psychology* **32**(5) 865–889.

Jong, B. A. de, K. T. Dirks, N. Gillespie. 2016. Trust and Team Performance: A Meta-Analysis of Main Effects, Moderators, and Covariates. *Journal of Applied Psychology* **101**(8) 1134–1150.

Komai, M., P. J. Grossman. 2009. Leadership and group size: An experiment. *Economics Letters* **105**(1) 20–22.

Larney, A., A. Rotella, P. Barclay. 2019. Stake Size Effects in Ultimatum Game and Dictator Game Offers: A Meta-Analysis. *Organizational Behavior and Human Decision Processes* **151** 61–72.

Ledyard, J. O. 1995. Public Goods: A Survey of Experimental Research. J. H. Kagel, A. E. Roth, eds. *The Handbook of Experimental Economics*. Princeton University Press, Princeton, 111–194.

McCannon, B. C. 2018. Leadership and Motivation for Public Goods Contributions. *Scottish Journal of Political Economy* **65**(1) 68–96.

Moxnes, E., E. van der Heijden. 2003. The Effect of Leadership in a Public Bad Experiment. *Journal of Conflict Resolution* **47**(6) 773–795.

Nosenzo, D., S. Quercia, M. Sefton. 2015. Cooperation in Small Groups: The Effect of Group Size. *Exp Econ* **18**(1) 4–14.

Rivas, M. F., M. Sutter. 2011. The Benefits of Voluntary Leadership in Experimental Public Goods Games. *Economics Letters* **112**(2) 176–178.

Sahin, S. G., C. Eckel, M. Komai. 2015. An Experimental Study of Leadership Institutions in Collective Action Games. *Journal of the Economic Science Association* **1**(1) 100–113.

Selten, R. 1967. Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments. H. Sauermann, ed. *Beiträge zur experimentellen Wirtschaftsforschung*. Mohr, Tübingen, 136–168.

Simons, T., H. Leroy, V. Collewaert, S. Masschelein. 2015. How Leader Alignment of Words and Deeds Affects Followers: A Meta-analysis of Behavioral Integrity Research. *Journal of Business Ethics* **132**(4) 831–844.

Teyssier, S. 2012. Inequity and Risk Aversion in Sequential Public Good Games. *Public Choice* **151**(1-2) 91–119.

Weimann, J., J. Brosig-Koch, T. Heinrich, H. Hennig-Schmidt, C. Keser. 2019. Public Good Provision by Large Groups – The Logic of Collective Action Revisited. *European Economic Review* **118** 348–363.

Yu, J., M. Kocher. 2020. Leading by Example in a Public Goods Experiment with Benefit Heterogeneity. *Working Paper*.

Zelmer, J. 2003. Linear Public Goods Experiments: A Meta-Analysis. *Experimental Economics* **6**(3) 299–310.

Zhang, Y., L. Zhang, G. Liu, J. Duan, S. Xu, M. W.-L. Cheung. 2019. How Does Ethical Leadership Impact Employee Organizational Citizenship Behavior? *Zeitschrift für Psychologie* **227**(1) 18–30.

APPENDIX

DESCRIPTIVE STATISTICS

Table A. 1: Descriptive Statistics for each study

Study and included treatments			Rel. contributions			Public Goods Environment			
	Structure	Observ.	All	Leader	Follower	G. Size	MPCR	Endow.	Periods
Centorrino & Concina (2013)									
X-Treatment	LbE	24	0.407 (0.173)	0.458 (0.206)	0.390 (0.187)	4	0.5	30	10
Dannenberg (2015)									
Ex-Base	Sim.	10	0.349 (0.250)	---	---	4	0.4	25	10
Ex-Leader	LbE	10	0.519 (0.220)	0.610 (0.227)	0.488 (0.230)	4	0.4	25	10
Drouvelis & Nosenzo (2013)									
No-ID	LbE	16	0.531 (0.249)	0.621 (0.218)	0.486 (0.275)	3	0.5	20	10
Eisenkopf & Kölpin (2021)[1]									
Small-Team (w/o Leader)	Sim	20	0.384 (.322)	---	---	3	0.5	100	20
Small-Team (w. Leader)	LbE	20	0.524 (0.290)	0.547 (0.288)	0.513 (0.295)	3	0.5	100	20
Large-Team (w/o Leader)	Sim	9	0.773 (0.166)	---	---	6	0.5	100	20
Large-Team (w. Leader)	LbE	11	0.727 (0.205)	0.846 (0.181)	0.704 (0.216)	6	0.5	100	20

[1] We pooled the data from the two Small-Team treatments (with and without observation of the other team's contribution).

Study and included treatments	**Structure**	**Observ.**	**Rel. contributions**			**Public Goods Environment**			
			All	**Leader**	**Follower**	**G. Size**	**MPCR**	**Endow.**	**Periods**
Eisenkopf (2020)									
Baseline	Sim.	16	0.301 (0.126)	---	---	3	0.5	100	20
Leading-by-Example	LbE	8	0.607 (0.295)	0.673 (0.257)	0.567 (0.302)	3	0.5	100	20
Eisenkopf & Walter (2021)[2]									
Baseline-IM/GM	Sim.	40	0.240 (0.145)	---	---	3	0.5	100	20
Lbe-IM/GM	LbE	40	0.502 (0.311)	0.526 (0.304)	0.490 (0.317)	3	0.5	100	20
Frackenpohl et al. (2016)									
Give-R	LbE	18	0.573 (0.331)	0.664 (0.329)	0.542 (0.338)	4	0.4	20	10
Gächter & Renner (2018)									
No-Leader-Treatment	Sim.	12	0.512 (0.179)	---	---	4	0.4	20	10
Leader-Treatment	LbE	12	0.482 (0.251)	0.543 (0.242)	0.462 (0.267)	4	0.4	20	10
Gürerk et al. (2018)									
Treatment P	Sim.	12	0.563 (0.231)	---	---	4	0.4	20	20
Treatment L	LbE	12	0.488 (0.286)	0.560 (0.282)	0.463 (0.293)	4	0.4	20	20

[2] We pooled the data from the individual and group monitoring treatments.

Study and included treatments	Structure	Observ.	Rel. contributions			Public Goods Environment			
			All	Leader	Follower	G. Size	MPCR	Endow.	Periods
Güth et al. (2007)									
C- Control	Sim.	14	0.402 (0.247)	---	---	4	0.4	25	16
Lf – Leader fixed	LbE	14	0.524 (0.294)	0.611 (0.284)	0.495 (0.301)	4	0.4	25	16
Moxnes & van der Heijden (2003)									
Control-Treatment	Sim.	12	0.137 (0.054)	---	---	5	0.4	20	10
Leader-Treatment	LbE	12	0.366 (0.142)	0.847 (0.027)	0.246 (0.176)	5	0.4	20	10
Rivas & Sutter (2011)									
Control	Sim.	14	0.40 (0.247)	---	---	4	0.4	25	16
Exogenous	LbE	14	0.350 (0.187)	0.473 (0.238)	0.310 (0.177)	4	0.4	25	16
Sahin et al. (2015) [3]									
Baseline-Treatment	Sim.	8	0.682 (0.067)	---	---	6	0.2	9	20
Exemplar-Treatment	LbE	14	0.731 (0.136)	0.855 (0.124)	0.706 (0.146)	6	0.2	9	20

[3] We excluded two groups in the Exemplar-Treatment because of an undefined role assignment.

Study and included treatments	Structure	Observ.	Rel. contributions All	Leader	Follower	Public Goods Environment G. Size	MPCR	Endow.	Periods
Yu & Kocher (2020)									
Baseline	Sim.	12	0.360 (0.096)	---	---	4	0.4/0.8	20	10
HBL & LBL	LbE	23	0.433 (0.199)	0.50 (0.278)	0.411 (0.209)	4	0.4/0.8	20	10

Note: The table shows the included studies and corresponding original treatment names. Moreover, it reports the independent observations as well as the average contributions in percent of the respective endowment. Standard deviations are reported in parentheses.

A.1 ROBUSTNESS CHECKS

This section provides robustness checks for our results given in the main part of our paper. In section 0 we replicate the regression based on groups with leaders, but exclude the groups from additional studies (see section IV.B). In section 0 we exclude certain studies that induce a change in the results.

Replication of regressions without the additional studies

This subsection includes robustness checks for the regression based on the groups with leaders. In the main part of our paper, we included additional groups from other studies that do not allow comparison between simultaneous contributions and Leading-by-Example (see Table 1 in section III.A as well as section IV.B). In the robustness tests presented here, we replicate the estimates from the main part but excluded the additional groups from the studies Centorrino and Concina (2013), Drouvelis and Nosenzo (2013) and Frackenpohl et al. (2016). This sub-section is ordered as follows: Table A. 2 replicates all six models from Table 4. Table A. 3 provides replications for the models IV-VI from Table 6. Last not least, Table A. 4 replicates all seven models from Table 8.

Table A. 2: Replication of Table 4: The impact of Leading-by-Example

Dep. Var.: Indiv. contribution in percent of the endowment	**Leading-by-Example**					
	Leader & Follower			**Only Followers**		
	I	**II**	**III**	**IV**	**V**	**VI**
Leader	0.088***	0.114***	0.113***			
	(0.013)	(0.011)	(0.011)			
Rel. Leader contribution				0.685***	0.718***	0.723***
				(0.025)	(0.026)	(0.024)
Group Size		0.090***	0.074**		0.010	-0.008
		(0.021)	(0.029)		(0.014)	(0.015)
Exchange rate (in €)		-0.359	-1.324		-1.554***	-2.446***
		(0.549)	(0.830)		(0.468)	(0.567)
MPCR		0.325*	0.653***		0.306*	0.799***
		(0.170)	(0.127)		(0.170)	(0.158)
Endowment		-0.001	-0.010*		-0.002***	-0.011***
		(0.001)	(0.005)		(0.001)	(0.003)
Total number of periods		0.014**	0.027***		0.021***	0.041***
		(0.006)	(0.006)		(0.005)	(0.005)
Leader not fixed	-0.174***	-0.122**	-0.170*	-0.056**	-0.026	-0.078**
	(0.040)	(0.051)	(0.089)	(0.028)	(0.033)	(0.038)
Fixed Effects for Studies	---	---	✓	---	---	✓
Constant	0.545***	-0.196	-0.149	0.085***	-0.318***	-0.471***
	(0.024)	(0.143)	(0.252)	(0.015)	(0.121)	(0.178)
Observations	12,232	12,232	12,232	9,096	9,096	9,096
Number of Groups	190	190	190	190	190	190
R-squared	0.034	0.093	0.114	0.402	0.447	0.478

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** $p<0.01$, ** $p<0.05$, * $p<0.1$*

Table A. 3: Replication of Table 6: Long-term impact of leadership (model IV-VI)

Dep. Var.: Indiv. contribution in percent of the endowment	**Leading-by-Example**		
	I	**II**	**III**
Leader	0.0973***	0.1108***	0.1083***
	(0.0207)	(0.0185)	(0.0183)
Leader*Period	-0.0010	0.0003	0.0005
	(0.0014)	(0.0013)	(0.0013)
Period	-0.0034*	-0.0068***	-0.0068***
	(0.0018)	(0.0015)	(0.0015)
Group Size		0.0900***	0.0740**
		(0.0214)	(0.0289)
Exchange rate (in €)		-0.3595	-0.0374
		(0.5488)	(0.4983)
MPCR		0.3251*	0.6530***
		(0.1696)	(0.1268)
Endowment		-0.0005	-0.0177***
		(0.0010)	(0.0056)
Total number of periods		0.0176***	0.0263***
		(0.0059)	(0.0067)
Leader not fixed	-0.1841***	-0.1218**	
	(0.0393)	(0.0506)	
Fixed Effects for Studies	---	---	✓
Constant	0.5778***	-0.1915	-0.1541
	(0.0258)	(0.1435)	(0.2504)
Observations	12,232	12,232	12,232
Number of Groups	190	190	190
R-squared	0.0361	0.1002	0.1219

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

Table A. 4: Replication of Table 8: The impact of group size in Leading-by-Example treatments

Dep. Var.: Indiv. contribution in percent of the endowment	Leading-by-Example						
	Leader			Follower			
	I	II	III	IV	V	VI	VII
Group Size	0.10***	0.11***	0.10***	0.06***	-0.013	0.041**	0.07***
	(0.016)	(0.019)	(0.027)	(0.017)	(0.011)	(0.020)	(0.025)
Rel. Leader contr. (RLC)					0.70***	0.98***	1.11***
					(0.027)	(0.099)	(0.092)
Group Size * RLC						-0.07***	-0.10***
						(0.026)	(0.025)
Exchange rate (in €)		1.64***	1.926**				-2.37***
		(0.491)	(0.852)				(0.570)
MPCR		0.63***	0.77***				0.80***
		(0.219)	(0.238)				(0.157)
Endowment		0.000	-0.004				-0.01***
		(0.001)	(0.006)				(0.003)
Total number of periods		0.004	0.009				0.04***
		(0.006)	(0.008)				(0.005)
Leader not fixed	-0.15***	-0.17***	-0.140	-0.16***	-0.058**	-0.068**	-0.079**
	(0.047)	(0.055)	(0.095)	(0.038)	(0.028)	(0.028)	(0.038)
Fixed Effects for Studies	---	---	✓	---	---	---	✓
Constant	0.22***	-0.214	-0.262	0.26***	0.13***	-0.070	-0.74***
	(0.074)	(0.161)	(0.302)	(0.073)	(0.039)	(0.068)	(0.190)
Observations	3,112	3,112	3,112	10,700	9,096	9,096	9,096
Number of Groups	190	190	190	190	190	190	190
R-squared	0.101	0.120	0.130	0.050	0.403	0.406	0.485

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

Replication of regressions without excluded studies

In this section, we replicate the regression given in our main part of the paper, but we exclude certain studies that induce a change in the result. This section is ordered as follows: Table A. 5 replicates Table 5. Table A. 6 provides replications of model I-III of Table 6 and Table A. 7 replicates model IV-VII of Table 8.

Table A. 5: Replication of Table 5: Standard deviations of individual contributions within groups

Dep. Var.: Standard deviations within groups	Excluded study: Eisenkopf & Walter (2021)			Güth et al. (2007)		
	Both Treatments Benchmark: Simultaneous					
	I	II	III	IV	V	VI
Leading-by-Example	-0.740	-0.372	-0.138	-0.017	-0.03***	-0.026**
	(1.383)	(0.956)	(0.984)	(0.012)	(0.010)	(0.010)
Group Size		3.07***	3.96***		0.04***	0.04***
		(0.990)	(1.172)		(0.010)	(0.012)
Exchange rate (in €)		-23.39**	-909.820		0.205	-32.461
		(10.622)	(927.683)		(0.210)	(32.944)
MPCR		6.36***	-0.000		0.130*	0.037
		(1.808)	(.)		(0.076)	(0.091)
Endowment		0.16***	-0.599		-0.000*	-0.033
		(0.017)	(1.118)		(0.000)	(0.033)
Total number of periods		-0.22***	7.059		-0.003*	0.271
		(0.053)	(7.748)		(0.002)	(0.283)
Fixed Effects for Studies	---	---	✓	---	---	✓
Constant	9.00***	-10.09**	-50.636	0.19***	0.075	-1.277
	(0.919)	(4.100)	(37.794)	(0.007)	(0.064)	(1.504)
Observations	4,526	4,526	4,526	5,678	5,678	5,678
Number of Groups	289	289	289	341	341	341
R-squared	0.000	0.100	0.110	0.003	0.134	0.139

OLS-Regression. Robust standard errors in parentheses clustered at Group-Level
**** p<0.01, ** p<0.05, * p<0.1*

Table A. 6: Replication of Table 6: Long-term impact of leadership (model I-III)

Dep. Var.: Indiv. contribution in percent of the endowment		**Excluded study: Eisenkopf & Walter (2021)** **Both Treatments** **Benchmark: Simultaneous**	
	I	**II**	**III**
Leading-by-Example (LbE)	0.106***	0.069**	0.060*
	(0.037)	(0.032)	(0.034)
LbE*Period	0.002	0.002	0.003
	(0.003)	(0.003)	(0.003)
Period	-0.005**	-0.011***	-0.011***
	(0.002)	(0.002)	(0.002)
Group Size		0.110***	0.092***
		(0.016)	(0.021)
Exchange rate (in €)		-1.399***	-13.143
		(0.373)	(92.071)
MPCR		0.233*	0.685***
		(0.126)	(0.103)
Endowment		-0.001**	0.029
		(0.001)	(0.032)
Total number of periods		0.023***	0.100
		(0.004)	(0.688)
Leader not fixed	-0.202***	-0.106**	-0.079
	(0.041)	(0.044)	(0.061)
Fixed Effects for Studies	---	---	✓
Constant	0.508***	-0.275**	-1.617
	(0.025)	(0.113)	(5.603)
Observations	18,744	18,744	18,744
Number of Groups	289	289	289
R-squared	0.036	0.152	0.170

OLS-Regression. Robust standard errors in parentheses clustered at Group-Level
**** p<0.01, ** p<0.05, * p<0.1*

Table A. 7: Replication of Table 8: The impact of group size in Leading-by-Example treatments (model IV-VII)

Dep. Var.: Indiv. contribution in percent of the endowment	**Excluded study: Sahin et al. (2015)**			
	Leading-by-Example Treatments Only Followers			
	I	**II**	**III**	**IV**
Group Size	0.0397 (0.0249)	-0.0301* (0.0170)	0.0333 (0.0227)	0.0757** (0.0317)
Rel. Leader contribution (RLC)		0.6921*** (0.0248)	1.0097*** (0.1314)	1.1233*** (0.1214)
RLC*Group Size			-0.0835** (0.0360)	-0.1083*** (0.0340)
Exchange (in €)				-2.3494*** (0.5725)
MPCR				0.7901*** (0.1567)
Endowment				0.0010 (0.0061)
Total number of periods				0.0305*** (0.0064)
Leader not fixed	-0.1427*** (0.0391)	-0.0459 (0.0284)	-0.0570** (0.0282)	-0.0824** (0.0391)
Fixed Effects for Studies	---	---	---	✓
Constant	0.3445*** (0.0974)	0.1920*** (0.0613)	-0.0425 (0.0788)	-0.9082*** (0.2291)
Observations	9,300	9,276	9,276	9,276
Number of Groups	234	234	234	234
R-squared	0.0234	0.4060	0.4098	0.4810

OLS-Regression. Robust standard errors in parentheses clustered at Group-Level
**** p<0.01, ** p<0.05, * p<0.1*

THURGAU INSTITUTE
OF ECONOMICS
at the University of Konstanz

Hauptstrasse 90
Postfach
CH-8280 Kreuzlingen 2

T +41 (0)71 677 05 10
F +41 (0)71 677 05 11

info@twi-kreuzlingen.ch
www.twi-kreuzlingen.ch