# Might There Be More Than a Grain of Truth in Self-Reported Effort Information? An Experimental Study

Özgür Gürerk[+], Thomas Lauer[++], Mark Pigors[++]

[+]RWTH Aachen University, [++]University of Cologne

**July 2014**

**Abstract:** We experimentally investigate the effects of teammates' self-reported effort information on the supervisor's performance pay allocation and on team performance. When reporting, teammates exaggerate their own efforts. However, they exaggerate less if the supervisor has the power to allocate individual payments at her own discretion than when the supervisor has to follow an exogenous allocation rule. The supervisor's ability to allocate incentive-compatible performance payments does not depend on whether she can monitor subordinates' true efforts or whether she receives self-reports. The exaggerations in self-reports have detrimental effects on team performance; these effects, however, are less pronounced if the supervisor allocates performance pay endogenously than when payments are distributed exogenously.

**Acknowledgements:** We thank the participants in the IMEBE 2013 in Madrid, and in the ESA World Meetings Zurich 2013, as well as seminar participants from Aachen, Erfurt, and Nottingham for useful comments. We also thank Anne Schielke for her research assistance in the conduction of the experiments. Financial support from the Deutsche Forschungsgemeinschaft, DFG, through GU 954/3 and LA 3372/1, is gratefully acknowledged.

**Addresses**

| Özgür Gürerk | Thomas Lauer | Mark Pigors |
|---|---|---|
| RWTH Aachen University | University of Cologne | University of Cologne |
| School of Business and Economics | Department of Economics | Department of Economics |
| Templergraben 64 | Albertus-Magnus-Platz | Albertus-Magnus-Platz |
| 52062 Aachen, Germany | 50923 Köln, Germany | 50923 Köln, Germany |
| ozgur.gurerk@rwth-aachen.de | thomas.lauer@uni-koeln.de | mark.pigors@uni-koeln.de |
| www.expecon.rwth-aachen.de | www.behavecon.uni-koeln.de | www.behavecon.uni-koeln.de |

## 1. Introduction

Organizations often pay their employees performance-contingent rewards, as many of Fortune 500 companies (Boyle 2001). Abundant evidence shows that pay for performance has a positive effect on employees' effort if it is closely related to individual performance (for an overview, see e.g., Milkovich and Newman 2008). In work teams producing a joint output, however, evaluating individual performance accurately is often difficult or even impossible. Consider, for instance, creative teams, where many ideas and solutions arise from sessions of brainstorming. In such teams, it is difficult to evaluate which teammate contributed what amount of effort and what the corresponding value is. One possible way to gather information about individual efforts is to ask the teammates for their self-assessments. Indeed, self-reports are frequently used in performance evaluations as additional data to confirm ratings from other sources, e.g., peer and supervisor evaluations. They are also an integral part of the evaluation in multi-source or 360-degree feedback processes (Campbell and Lee 1988, Chen and Kemp 2012).

The empirical evidence from the organizational literature is inconclusive about the effectiveness of self-reports (see Dunning et al. 2004, for a review) and poses the question of whether self-reports really do add any value to an effective performance evaluation. Furthermore, it is not clear as to how self-reports ultimately affect a team's performance (Shore and Tashchian 2002). Previous studies from the organizational literature based mainly on questionnaires, show that teammates tend to overestimate their own efforts or even deliberately lie about their actual efforts (Campbell and Lee 1988). Self-reports may also have effects on supervisors' rating behavior and cause them to rate subordinates leniently (see, e.g., Prendergast 1999). Recent experimental studies show that a substantial amount of people do indeed lie if self-reports have monetary consequences and cannot be verified (Fischbacher and Föllmi-Heusi, 2013). On the other hand, many people show a lying aversion that varies to different extents, as experimental studies also show (see, e.g., Erat and Gneezy 2012, Gibson et al. 2013, Gneezy et al. 2013). A recent representative survey confirms the existence of lying costs in the population (Abeler et al. 2014). These findings lead us to consider whether there might be more than a grain of truth in self-reports.

To shed more light on self-reports' truthfulness as well as their impact on team performance, in this study, we investigate the effect of self-reports in a controlled laboratory setting. Laboratory experiments allow the unveiling of casual relationships between the variables of interest. Since previous non-laboratory studies lacked the possibility to observe true efforts, they could not quantify the frequency and magnitude of the exaggerations in self-reported effort information. To understand the effects and the value of self-reports, however, it is important to know the direction and amount of deviances from true (actual) efforts. We designed this experimental study to detect truthfulness in self-reports in a quantifiable way. To our best knowledge, this is the first experimental study concerning

the effects of self-reports on team performance. Additionally, our study contributes to the growing experimental literature on lying and other unethical behavior in the workplace.

We investigate four research questions. (1) Do teammates exaggerate their efforts when the allocation of performance pay is set exogenously? (2) Is teammates' reporting behavior different when a supervisor endogenously decides on performance pay compared to a situation when the allocation rule is exogenous (and the supervisor is only an observer)? (3) Are supervisors able to allocate incentive-compatible payments when receiving self-reports as effectively as when they are able to monitor true individual efforts? (4) Finally, what is the ultimate effect of self-reports on team performance?

We find: (1) Teammates systematically exaggerate their efforts when reporting. (2) They exaggerate less when the supervisor has the power to allocate performance pay discretionally than when the payments are allocated exogenously. (3) The supervisor's allocation behavior regarding performance pay does not change, irrespective of whether the supervisor observes true individual efforts or self-reports. (4) The exaggerations in self-reports have detrimental effects on team performance. These effects are less pronounced if the supervisor allocates performance pays endogenously than when the payments are distributed exogenously.

In the following section, we propose a simple model of team production. Section 3 explains our treatments and the experimental procedure. In section 4, we state our hypotheses in light of our model and previous experimental results. Section 5 presents the results, and section 6 concludes.

## 2. A simple model of team production

We consider a work team pretty much in the spirit of Alchian and Demsetz (1972), consisting of three productive *teammates* and a *supervisor*. To capture the tension between individual and collective interest in a work team, we utilize a slightly modified public goods game. Each teammate *i* is endowed with $E_T$ tokens and decides how many tokens to contribute to the team project, i.e., his or her effort choice, $e_i$, with $0 \leq e_i \leq E_T$. The not contributed tokens, $E_T - e_i$, go into the own private account of teammate $i$. Each token contributed to the team project is multiplied by the productivity factor $M$, with $M/3 < 1 < M$. The team output, $R$, is the sum of the efforts $W = \sum_{i=1}^{3} e_i$, multiplied by the productivity factor $M$, i.e., $R = MW$.

To conduct a committed observation and action, the supervisor receives a compensation that partly depends on the team output. The team output is allocated among the three teammates and the supervisor as follows: $R$ is divided into three parts: $\alpha R$, $\beta R$ and $\gamma R$, with $\alpha + \beta + \gamma = 1$. The first part, $\alpha R$, describes a variable team payment that depends on the overall team performance, and is equally distributed to teammates. Hence, independent of the individual effort, each teammate receives $\frac{\alpha}{3}R$. The second part, $\beta R$, represents the individual performance-contingent pay. Each teammate receives a share, $s_i$ of $\beta R$, with $\sum_{i=1}^{3} s_i = 1$. The last part, $\gamma R$, is used to compensate the supervisor.

Additionally, the supervisor receives a fixed payment of $E_S$ tokens. Since the compensation of the supervisor depends partly on the team output, the supervisor has an interest in the team output being as large as possible. Figure 1 visualizes our model.

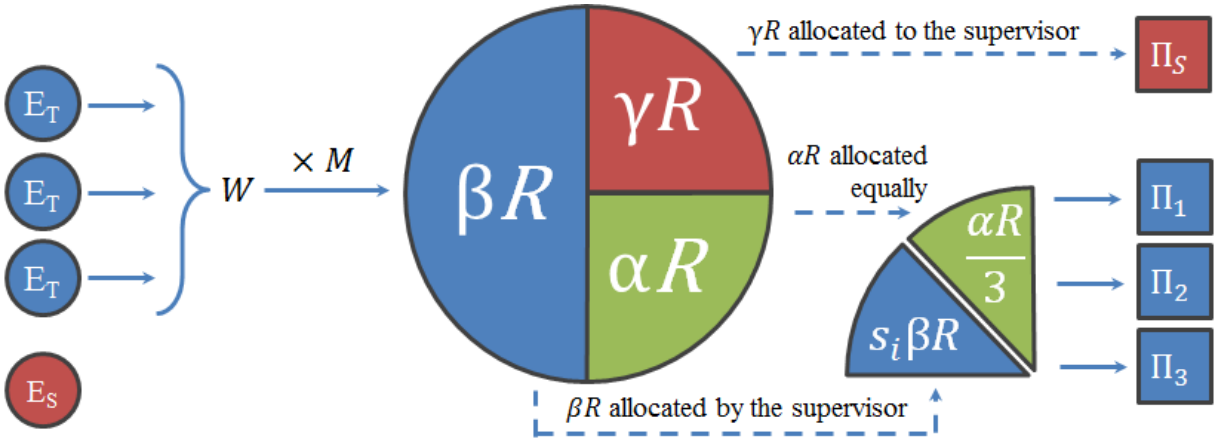



Figure 1: Basic model of efforts and payoff in our teamwork setting

Equations 1 and 2 state the payoffs for a teammate and the supervisor respectively:

$$\Pi_i = E_T - e_i + \frac{1}{3} \cdot \alpha R + s_i \cdot \beta R, \qquad i = 1,2,3 \tag{1}$$

$$\Pi_S = E_S + \gamma R. \tag{2}$$

In our design, we set the parameters as follows: $M = 2$, $\beta = 1/2$ , $\alpha = \gamma = 1/4$, $E_T = 20$, and $E_S = 10$.[1] With these parameters, we can re-formulate the payoffs as follows:

$$\Pi_i = 20 - e_i + \frac{1}{6} \cdot \sum_{j=1}^{3} e_j + s_i \cdot \sum_{j=1}^{3} e_j \,, i = 1,2,3 \tag{3}$$

$$\Pi_S = 10 + \tfrac{1}{2} \cdot \textstyle\sum_{j=1}^{3} e_j. \tag{4}$$

From (4), it is easy to see that the supervisor's payoff is maximized if all teammates contribute their full endowment to the team project, i.e., if each teammate exerts full effort. The supervisor can induce

1 We choose these parameter values for two reasons. First, they allow us to keep the marginal per capita return (MPCR) close to the value of most previous experiments on public good provision. In our setup, the MPCR amounts to $\frac{M}{3} = \frac{2}{3}$. Second, our set of parameters eases the task of the supervisor. Since the endogenously distributed part of the team-output equals the sum of efforts, the appropriate (incentive-compatible) proportional share is equal to each teammate's effort.

maximum efforts by distributing the variable part, $\beta R$, of the team output proportional[2] to the teammates' relative efforts, i.e., by choosing the shares such that $s_i = \frac{e_i}{\mathrm{W}}$. In this case, a teammate's payoff is

$$\Pi_i = 20 + \frac{1}{6} \cdot \sum_{j=1}^{3} e_j. \tag{5}$$

As can be seen in (5), full effort is a strictly dominant strategy for every teammate. Hence, assuming full information, in the equilibrium (of the stage game), the supervisor should allocate a performance pay share to each teammate proportional to his or her effort rate (compared to the team's total effort). A supervisor who simply distributes equally among the teammates, $s_i = 1/3$, would change each teammate's profit in (4) to $\Pi_i = 20 - e_i + \frac{1}{2} \cdot \sum_{j=1}^{3} e_j$. This change transfers the strategic situation in a public good game with a marginal per capita rate of ½ and a unique Nash equilibrium in which all teammates contribute zero. In public good experiments, however, subjects do contribute positive amounts (Chaudhuri 2011). Thus, under equal distribution of $\beta R$, we expect effort levels comparable to previous public good experiments. The positive externality of contributions on the supervisor's payoff in our setting, however, might have an effect on teammates' efforts. Engel and Rockenbach (2010), for instance, find that in a public goods experiment with a non-contributing bystander, efforts were lower than in the standard setting without a bystander.

## 3. Experimental Design and Procedures

The experiment consists of 20 rounds, in which subjects play the game explained above. Before the first round, subjects are randomly allocated into groups of four. In each group, one subject is randomly assigned as the supervisor and three subject as teammates.[3] The subjects' roles as well as the group matching remain unchanged during the experiment. Hence, each group constitutes an independent observation.

---

2 The proportional distribution scheme is not the only scheme that incentivizes the teammates to high effort levels, e.g., "the winner takes it all" is another incentive-compatible payment scheme. Yet we consider the proportional allocation to be the most obvious one under the given parameters. The results in section 5 show that the great majority of supervisors stick to the proportional distribution rule.

3 We are aware of the relevance of legitimating the supervisor by implementing some sort of selection process. However, considering different forms of appointment would have added another dimension to our design. Therefore, we use a random selection process, from which we expect the smallest impact on actual efforts and allocation decisions.

### 3.1 Treatments

In order to investigate our research questions, we apply a 2x2 full fractional design. First, we vary the information that the supervisor receives. The supervisor is informed about either the true individual efforts or the teammates' self-reports. Second, we manipulate how the performance pay $\beta R$ is distributed. In the exogenous treatments, it is distributed exogenously and proportionally to the actual or reported amounts. In the endogenous treatments, the supervisor allocates $\beta R$ at her own discretion. This means, in the exogenous treatments, that the supervisor is an observer with no distributive power, and is paid according to the supervisor's payoff function (cf. equation (4)). In the exogenous treatments, the performance pay shares $s_i$ are determined proportional; either according to the self-reports or to the actual efforts. In the treatment EXOGEN, the supervisor observes the true individual efforts, $e_i$, so the proportional distribution mechanism implies: $s_i = \frac{e_i}{W}$. In the treatment EXOGEN-SR, the $s_i$, are determined according to the self-reports, i.e., $s_i = \frac{r_i}{\sum_{i=1}^{3} r_i}$, with $r$ being the self-report of player $i$, and $\sum_{i=1}^{3} r_i$ being the sum of the reported efforts.

In the endogenous treatments, the supervisor discretionally decides about $s_i$. In ENDOGEN, before deciding, the supervisor is informed about the true efforts, while in ENDOGEN-SR the supervisor receives the self-reports. Table 1 displays our 2x2 experimental design.

| | Supervisor's information based on the… | |
|---|---|---|
| Distribution mechanism | Self-report | Actual effort |
| Endogenous allocation | ENDOGEN-SR | ENDOGEN |
| Exogenous allocation | EXOGEN-SR | EXOGEN |

Table 1: Treatment design

In addition to the four treatments introduced above, we conduct a "baseline" treatment named EQUAL, which uses a simple equal split as an exogenous distribution mechanism, i.e., each teammate's performance pay share equals to $s_i = \frac{1}{3}$. This transforms the game into a standard public-goods game, with the supervisor as bystander, since $\alpha R$ as well as $\beta R$ is allocated equally to the teammates.

### 3.2 The Timing of a Period

A period starts with the elicitation of subjects' beliefs about the other teammates' effort decisions. After the belief elicitation, each teammate chooses his or her own effort decision. Then, teammates are informed about the actual (true) efforts of fellow teammates. Following that, except in EQUAL, each

teammate states his or her own preferred distribution of the individual performance pay, i.e., each teammate states a distribution of the partial team output $\beta R$. Following this, depending on the treatment, the supervisor is either informed about the individual self-reports or the actual efforts. In the endogenous treatments, teammates additionally state their beliefs about the supervisor's distribution decision. In all treatments, the supervisor is informed about the true value of the team's total effort. After receiving the information, in the endogenous treatments, the supervisor decides about the individual shares, $s_i$, while in the exogenous treatments, the shares are allocated exogenously as described above. At the end of each period, each teammate receives feedback about his or her own payoff in the current period. Table 2 summarizes the timing of a period.

| **Teammate** | **Supervisor** |
|---|---|
| …states beliefs about fellow teammates' efforts | …states belief about teammates' efforts |
| …decides on actual own effort[4] | |
| …receives feedback about actual individual efforts (in all treatments) | …receives feedback about the actual total effort |
| …states preferences on allocation of $\beta R$ | |
| …submits self-report (in SR-treatments) | …receives feedback, depending on treatment, about actual efforts or self-reports |
| …states belief about supervisor's allocation (in endogenous treatments) | …allocates the shares, $s_i$[5], in exogenous treatments, the shares are allocated proportional to efforts/reports |
| …receive feedback about payoffs | |

Table 2: Timing of a period

### 3.3 Procedures

The experiment was conducted in the Cologne Laboratory for Economic Research (CLER). In total, 156 students from different disciplines participated. We collected eight independent observations (groups) per treatment for EXOGEN, EXOGEN-SR, ENDOGEN, and ENDOGEN-SR, and seven for EQUAL.[6] On their arrival, we informed subjects about the experimental procedure.[7] The experiment

4 Subjects had to choose their effort level from the set {0, 5, 10, 15, 20}. This restricted set simplifies the task for the supervisor and makes exaggerations meaningful.

5 To simplify the allocation decision, the supervisor enters the individual payoff, $s_i\beta R$, directly instead of calculating $s_i$ for each teammate.

6 Since some subjects did not show up, we had to run this session with seven groups.

was programmed with z-Tree (Fischbacher, 2007), and the subjects were recruited via ORSEE (Greiner, 2004). Each experimental session lasted about 90 minutes, and subjects earned about 18 Euros on average.

## 4. Predictions

### 4.1 Do teammates exaggerate in self-reports if the allocation of performance pay is exogenous?

The accuracy of self-reports in teamwork situations is subject to at least two problems (Campbell and Lee 1988). First, individuals may not be able to evaluate and report their own efforts objectively and reliably (overestimation). Second, they may deliberately misreport their efforts. In our study, we rule out by design the existence of (unintended) overestimation, since an individual is always able to evaluate his or her own effort perfectly. Hence, in our setting, if a teammate misreports, he or she must do so deliberately.

To answer the question of whether teammates do indeed lie about actual efforts and exaggerate, we utilize our treatment, EXOGEN-SR, in which lying is very advantageous. As explained before, in EXOGEN-SR, the performance payments are allocated exogenously (and proportionally) according to self-reports. Since the exogenous proportional allocation is incentive compatible in the sense the higher the report the higher the performance pay, for the teammates it maximizes the incentives to lie. For a teammate, it is strictly dominant to report maximum effort. On the other hand, for each teammate, it is rational to set one's actual effort to zero. Considering these two motives, we state our first hypothesis as follows:

**Hypothesis 1** In EXOGEN-SR, we expect the average level of self-reports to be clearly higher than the actual efforts.

### 4.2 Do teammates exaggerate/lie to a lesser extent if the supervisor endogenously decides on the performance pay shares?

Next, we ask whether teammates' reporting behavior might be different when the supervisor endogenously decides about the performance pay shares. To investigate this question, we contrast the reporting behavior from the EXOGEN-SR treatment explained above with ENDOGEN-SR in which the supervisor allocates the performance payments at her own discretion after receiving self-reports.

Should we expect a different behavior in ENDOGEN-SR how teammates' choose their reports and actual effort levels, compared to EXOGEN-SR? The answer depends on teammates' expectations of the supervisor's distribution in ENDOGEN-SR. If teammates believe that the supervisor in ENDOGEN-SR will deviate from the proportional distribution, e.g., if she decides to choose an

---

7 See the Appendix for a translation of the instructions. Original instructions are written in German. They are available upon request from the authors.

egalitarian distribution, then the teammates might have a motive to report less than maximum effort. In this case, in fact, teammates could report any amount. If, however, teammates expect the supervisor to stick to the proportional allocation, then they have an incentive to report maximum efforts – exactly as in EXOGEN-SR.

Indeed, the supervisor in ENDOGEN-SR has no reason to apply an other distribution than the proportional one, since all other distributions would decrease her expected payoff compared to the proportional distribution (see the footnote for a rationale for this argument).[8] Hence, the teammates in ENDOGEN-SR may correctly anticipate the proportional distribution and exaggerate their own efforts maximally.

**Hypothesis 2a:** We expect no differences, either in the average level of reported efforts or in the average level of actual efforts between the treatments EXOGEN-SR and ENDOGEN-SR. Thus, we do not expect a different level of lying (exaggeration) in ENDOGEN-SR and in EXOGEN-SR.

Recent experimental studies show that some people have an aversion to lying that varies to different extents as laboratory studies (Vanberg 2011, Gneezy et al. 2013, Conrads et al. 2014) as well as a study with a representative sample shows (Abeler et al. 2014). These "lying costs" could be more salient in the ENDOGEN-SR treatment, i.e., when the supervisor deliberately decides on the performance pay shares (than compared to the EXOGEN-SR treatment in which the supervisor does not make a decision). Subjects might have a stronger reluctance to lie to a supervisor with distributive power as in ENDOGEN-SR, than to a "lame-duck", observing-only supervisor, as in EXOGEN-SR. Consequently, in ENDOGEN-SR, teammates may not report full efforts. Thus, we formulate an alternative hypothesis 2b:

**Hypothesis 2b:** Teammates exaggerate less in ENDOGEN-SR than in EXOGEN-SR.

---

8 Why exactly should the supervisor choose a proportional distribution of the performance pay βR? Assume that the supervisor observes that the sum of the reported efforts is higher than the actual total contribution. Since there is no possibility to identify the exact amount of individual deviations in reports from the actual efforts, the supervisor is not able to discriminate between the teammates who lied and those who stated their true effort. By design, the supervisor is also not able to sanction the team as a whole by withholding payments, because the supervisor must allocate the entire amount of $\beta R$. An egalitarian allocation of $\beta R$ is also not feasible, since the equal allocation would transform the game into a public goods game. In public goods games, it is very likely that the team output will be low or decrease over time, thus lowering the supervisor's payoff. Hence, for the supervisor, with a very high probability, deviating from the proportional distribution would lead to worse outcomes than if she applied the proportional allocation.

### 4.3 Are supervisors able to allocate incentive compatible payments when receiving self-reports as effective as when they are informed about the true individual efforts?

As explained in the game analysis, to maximize own payoff, the supervisor should allocate performance pay proportional to individual efforts. To conduct an efficient allocation, the supervisor must know the true individual efforts. In the ENDOGEN treatment, we provide the supervisor with true information about the individual efforts. Hence, in ENDOGEN, the supervisor is able to motivate the teammates to contribute by setting $s_i \beta R = e_i$.

In the ENDOGEN-SR treatment, the supervisor receives individual self-reports as feedback. If teammates exaggerate their own efforts, the supervisor is faced with uncertainty. This uncertainty might render her allocation of performance pay ineffective for two reasons. First, if teammates exaggerate their efforts, then the proportional allocation of payments is not incentive-compatible. Second, since the supervisor is not able to monitor the actual efforts, she cannot condition the payments on efforts in any other incentive compatible way. Hence, with self-reports, we expect the supervisor's allocation of performance payments to be less effective in using incentive compatible performance pay than with true effort information. This ineffectiveness might lead to lower actual efforts in ENDOGEN-SR than in EXOGEN-SR, in which the supervisor is able to monitor true efforts.

**Hypothesis 3:** When receiving self-reports, supervisors are less effective in using incentive-compatible performance pay than when they receive true effort information.

## 5. Results

### 5.1 Subjects exaggerate self-reports when performance pay is allocated exogenously

We start our analysis by looking at the reporting behavior in EXOGEN-SR. Remember, the proportional allocation of performance pay according to self-reports in the EXOGEN-SR treatment provides a strong incentive to exaggerate. As Panel a) of Figure 2 shows, teammates in EXOGEN-SR do indeed exaggerate their efforts when reporting.[9] Actual efforts amount on average to 13.4, while self-reports amount to 16.5. The difference between the reported and the actual efforts is of considerable magnitude (3.1) and statistically significant (Pitman-Fisher permutation test[10], one-sided, $p$ = 0. 008). These numbers provide clear support for our hypothesis 1.

---

9 There is no case, in which a teammate exaggerated her effort when the actual sum of efforts in the group was zero, i.e., there is no case, in which a supervisor was able to detect an individual exaggeration for sure.

10 We decided to use the Fisher-Pitman test instead of the Wilcoxon rank sum test. While both tests are permutation tests, the Wilcoxon test is based on ranks and therefore ignores a substantial part of

**Result 1** If performance payments are allocated exogenously according to the self-reported effort information, then self-reported efforts are significantly higher than the actual efforts.

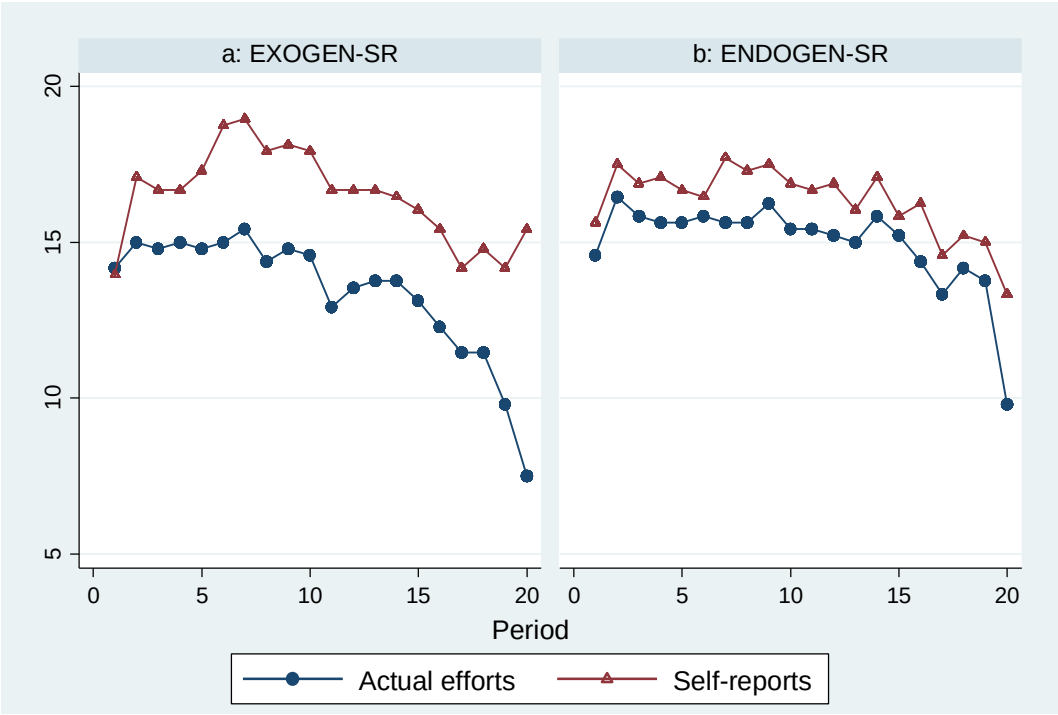



Figure 2: Actual and reported efforts

### 5.2 Subjects exaggerate self-reports less when performance pay is allocated endogenously by the supervisor

Do teammates report differently when the supervisor discretionarily decides on the allocation of performance pay, as in ENDOGEN-SR? We see in Figure 2, panel b, the pattern of reporting behavior is similar to EXOGEN-SR, as predicted in our hypothesis H2a. As in EXOGEN-SR, in the ENDOGEN-SR treatment, self-reported efforts (16.3) are significantly higher than the actual efforts (14.9, PFP test, one-sided, $p$ = 0.008). The magnitude of the exaggeration in ENDOGEN-SR, however, is clearly smaller than in EXOGEN-SR (1.4 versus 3.1). Teammates, on average, exaggerate to a lesser extent if the supervisor has allocation power than when she is just an observer. However, based on non-parametrical tests, neither the averages of reports nor the means of efforts, nor the average extent (amount) of exaggerations differ significantly between EXOGEN-SR and ENDOGEN-SR.

---

the information in the sample data. The Pitman-Fisher permutation test (PFP test) uses the more powerful approach based on the original sample values without transformation (cf. Kaiser 2007).

**Result 2a** The overall pattern of reporting behavior, the average level of actual efforts as well as of the reports are not statistically different between ENDOGEN-SR and EXOGEN-SR.

In the following, a closer look at the exaggerations will uncover some substantial differences in reporting behavior between ENDOGEN-SR and EXOGEN-SR. Panel a) of Figure 3 presents the frequency of exaggerations as the ratio of the actual untruthful self-reports to the number of cases with exaggeration possibilities, i.e., to all cases in which actual efforts are below 20. There are 218 such possible exaggeration cases in EXOGEN-SR and 197 in ENDOGEN-SR. On average, teammates exaggerate more often in EXOGEN-SR than in ENDOGEN-SR (62.8% versus 47.7%). This difference is significant according to Fisher's exact test ($p = 0.002$).[11]

Moreover, if we consider only those teammates who lied and not those who reported truthfully, then we see that teammates in EXOGEN-SR inflate their efforts considerably more than the teammates in ENDOGEN-SR. Panel b) of Figure 3 presents the average amount of exaggeration for the cases in which the reported effort is strictly higher than the actual effort.[12] The average deviation in these cases amounts to 13.2 in EXOGEN-SR, whereas it is 8.2 in ENDOGEN-SR. The difference is weakly significant according to the PFP test ($p = 0.076$). A random effects panel regression reveals a highly significant negative effect of the endogenously deciding supervisor on teammates' extent of exaggeration.[13] The coefficient of the "endogenous supervisor" is -3.216, the robust standard error is 1.55 and the p-value is 0.042. This fact confirms the difference in reporting behavior between the EXOGEN-SR and ENDOGEN-SR.

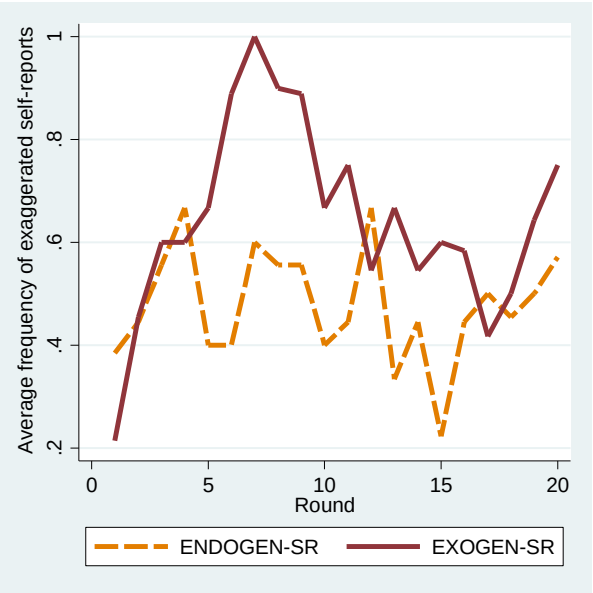
**Result 2b** In EXOGEN-SR, teammates exaggerate more often and stronger than in ENDOGEN-SR.

---

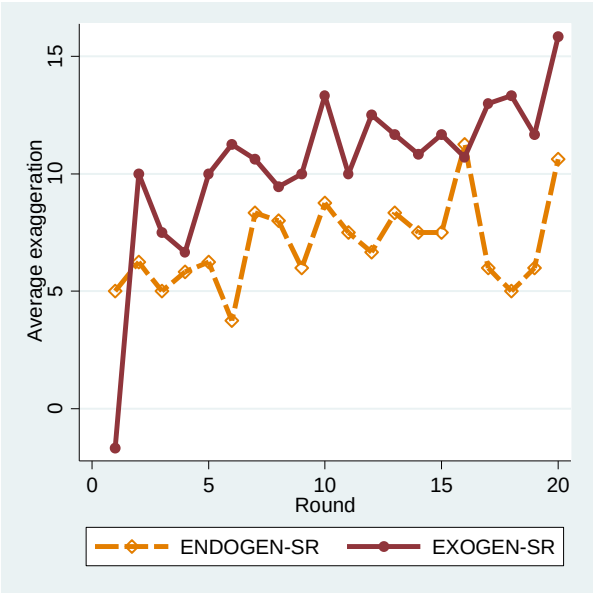
11 We are aware of the fact that Fisher's exact test does not consider clustering within groups. However, since we are primarily interested in the overall difference between the two treatments, we decided to use this test.

12 We exclude the cases, in which subjects reported lower amounts than their actual efforts. There are two such cases in EXOGEN-SR and 12 in ENDOGEN-SR. If we include these cases, the frequency of untruthful self-reports increases to 63.8% in EXOGEN-SR, and to 53.8% in ENDOGEN-SR.

13 We used a random effects panel regression for censored data with robust standard errors clustered for groups and "contribution" as additional independent variable.




Panel a)




Panel b)

Figure 3

Taken both results 2a and 2b together, we find some substantial support for the hypothesis 2b that teammates lie to a less extent in ENDOGEN-SR than in EXOGEN-SR.

### 5.3 Supervisors are able to allocate incentive compatible payments when receiving self-reports as effectively as when they are informed about actual individual efforts

To investigate whether supervisors are able to allocate performance pay in an incentive-compatible way, we conducted the treatment ENDOGEN, in which the supervisor is informed about teammates' true efforts. Indeed, in ENDOGEN, in 97.7% of the cases, supervisors allocate incentive compatible rewards. In the ENDOGEN-SR treatment, the respective number is 95.5% with respect to reports (89.7% with respect to actual efforts)[14]. The difference between both treatments is statistically not significant.

---

14 We measure a supervisor's performance as the deviation from the proportional distribution benchmark. As mentioned above, the proportional distribution is not the only incentive-compatible allocation rule. However, the results from ENDOGEN show that almost all chosen distributions were proportional. Formally, the supervisors' quality is defined by the normed geometric distance from the perfect proportional allocation $1 - \frac{1}{\sqrt{2}}\sum_{i=1}^{3}(s_i - s_i^*)^2$ where $s_i^*$ is member $i$'s share in the proportional distribution based on contributions $\left(s_i^* = \frac{e_i}{W}\right)$ or on reported contributions in ENDOGEN-SR.

**Result 3** Supervisors in ENDOGEN-SR, who receive self-reports, are able to apply an incentive compatible performance pay allocation as effectively as supervisors in ENDOGEN who are informed about actual efforts.

### 5.4 The effect of self-reports on team performance

How do self-reports ultimately affect team performance? The answer depends on to which setting we compare. We could compare self-report treatments to an ideal setting of perfect monitoring and perfectly incentivizing performance pay allocations. In the EXOGEN treatment, we set up such an ideal setting. As Figure 4 shows compared to EXOGEN, both self-report treatments have detrimental effects on team performance. The clear difference in average efforts between EXOGEN (19.3 on average) and EXOGEN-SR is statistically significant (13.4, PFP test, one-sided $p = 0.033$). Compared to the other extreme situation of a standard public goods setting as in EQUAL, in which average effort amounts to 11.1, self-report treatments perform better, although the differences are not significant according to PFP test ($p = 0.236$ for EXOGEN-SR and $p = 0.117$ for ENDOGEN-SR).

**Result 4a** Compared to an ideal situation in EXOGEN, self-reports have detrimental effects.

The comparison of average efforts in ENDOGEN and ENDOGEN-SR shows that there is no significant detrimental effect of self-reports when the supervisor decides endogenously. Although the average efforts tend to be smaller in ENDOGEN-SR (14.9) than in ENDOGEN (18.1), this difference is statistically not significant (PFP test, one-sided $p = 0.107$).

**Result 4b** Self-reports have no significant detrimental effect on team performance if the supervisor has distributive power.

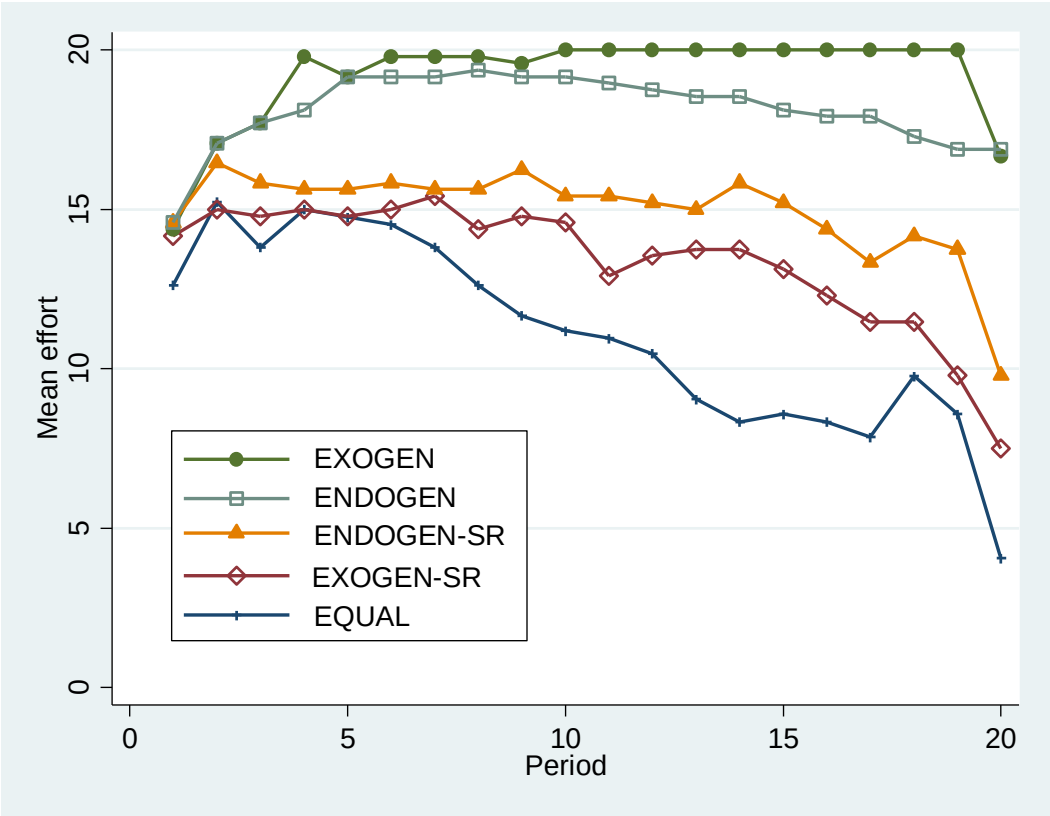



Figure 4: Average team performance

The positive effect of having a supervisor with distributive power under self-reports becomes obvious when we focus on the average distances between EXOGEN and EXOGEN-SR, and between ENDOGEN and ENDOGEN-SR respectively. Whereas self-reports reduce the average efforts by about 5.94 tokens if the distribution of performance pay is exogenous (EXOGEN minus EXOGEN-SR), the negative effect of self-reports is only 3.18 tokens between both endogenous treatments (ENDOGEN minus ENDOGEN-SR).

**Result 4c** The detrimental effects of self-reports on team performance are smaller when the supervisor has distributive power than when the performance pay shares are allocated exogenously.

### 5.5 Regression analysis

In a set of random effects panel regression (see Table 3), we can confirm all of the results that we have presented so far. We assume that the following model(s) explains the variance in teammates' efforts. The first explanatory variable that we add to all three variants of the model is a subject's belief about his or her teammates' average effort in the next period (*Belief about others' average efforts*). Similar to Fischbacher & Gächter (2010) we assume that subjects update this belief depending on the received feedback in a first step, before they use it to determine their next period's effort. The second

explanatory variable is a measure for *supervisor performance*[15] that we construct as "1 - the average absolute distance" between the chosen share for the teammate (by the supervisor) and the share the teammate would receive if the shares would be allocated perfect proportional to the actual efforts. Since the teammates learn the supervisor's allocation decision only at the end of each round, we include the first lag ($t_{-1}$) of the supervisor performance measure. The dummy variable *Equal distribution* is 1 in the EQUAL treatment in which all teammates receive the same performance pay share irrespective of their relative effort. The second dummy *Self-report* is set to 1 in ENDOGEN-SR and EXOGEN-SR. The dummy *Endogenous supervisor* is 1 for the treatments in which the supervisor allocates at her own discretion (ENDOGEN and ENDOGEN-SR). Finally, we include an interaction term between the self-report dummy and the endogenous supervisor dummy that is set to one only in ENDOGEN-SR. The results of the random effects panel regressions with robust standard errors clustered on groups are presented in Table 3.

The *belief about others' average effort* have a positive effect on a teammate's own effort in each of the three variants of the model, i.e., the higher a teammate's expectation of others' average effort , the more he or she is willing to contribute. The regression also shows a large positive effect of *Supervisor performance in* $t_{-1}$ on efforts. The overall effect of supervisor's ability to match the perfect proportional share is in line with the results obtained with non-parametrical tests for the comparison between the ENDOGEN and EXOGEN treatment.

The negative coefficient for *Equal distribution* (see model I in Table 3) confirms the positive effect of performance based payments in the exogenous and endogenous treatments compared to the EQUAL treatment. The dummies for *Self-reports* confirm the results of the nonparametric tests (see model II in Table 3). If the payment is mainly based on self-reports as in the two self-report treatments, the teammate's effort decreases by about 1.771. The coefficient for the interaction between endogenous supervisor and self-reports (*endogenous supervisor x self-reports*) shows the positive effect that a supervisor with distributive power on efforts, i.e., the detrimental effects of self-reports are much less pronounced if there is an endogenous supervisor (see model III in Table 3).

---

15 We measure the supervisor's performance as the deviation from the proportional distribution benchmark. As mentioned above, the proportional distribution is not the only incentive compatible allocation rule. However, the results from ENDOGEN show that almost all chosen distributions were proportional. Formally, the supervisor's quality is defined by the normed geometric distance from the perfect proportional allocation $1-\frac{1}{\sqrt{2}}\sum_{i=1}^{3}(s_i-s_i^*)^2$ where $s_i^*$ is teammate $i$'s share in the proportional distribution based on actual efforts $\left(s_i^*=\frac{e_i}{W}\right)$ or on reported efforts in ENDOGEN-SR.

| Dependent variable: Effort in t | I | II | III |
|---|---|---|---|
| Belief about others' average efforts | 1.066***<br>(0.054) | 1.057***<br>(0.059) | 1.051***<br>(0.064) |
| Supervisor performance in $t_{-1}$ | 3.179***<br>(0.845) | 2.639**<br>(1.256) | 2.937**<br>(1.400) |
| Equal distribution | -2.121**<br>(1.056) | | |
| Self-reports | | -1.771**<br>(0.708) | -2.429**<br>(1.079) |
| Endogenous supervisor | | | -0.375<br>(0.303) |
| Endogenous supervisor x Self-reports | | | 1.306<br>(1.294) |
| Constant | 1.907***<br>(0.707) | 3.413**<br>(1.607) | 3.373**<br>(1.461) |
| *N* | 2223 | 1824 | 1824 |
| $R^2$ | 0.672 | 0.691 | 0.693 |

Table 3: Random effects panel regression on effort, robust standard errors clustered on groups

## 6. Conclusion

We experimentally complement the finding from the organizational behavior literature that one has to be cautious when implementing and evaluating self-reports in performance appraisals. Compared to an ideal but rather unrealistic situation of full information, the ultimate effect of self-reports on team performance is negative. The detrimental effects of self-reports, however, are less pronounced when a supervisor has distributive power to allocate performance pay than under an exogenous allocation mechanism where the supervisor is only a bystander.

Teammates exaggerate self-reports strongly when the allocation mechanism of performance pay is exogenous and known ex-ante. If the supervisor possesses the power to allocate performance pay at her own discretion and ex-post, self-reports are less exaggerated. Interestingly, the uncertainty stemming from the fact that the supervisors ex-post decide on performance pay shares at their own discretion seem to increase efforts and decrease the amount of exaggerations. Translating this result as a practical implication for reality would mean the following: In situations where a supervisor must rely (at least to some extent) on self-reported effort information, it might be better for the management not

to announce hard-wired performance pay ex-ante. Management should rather, keep the possibility open of determining/adjusting performance pay until observed the joint outcome and the (possibly exaggerated) self-reports. The results point to a possible “cost of transparency” when a performance pay scheme is hard-wired.

Of course, as always, practical implications of experimental results should be interpreted with caution. The transferability of this study’s results to reality might be limited by the fact that we have rather a strong test for self-reporting. In reality, the detection of false self-reports might be easier. Hence, our findings may overestimate the (detrimental) effects of self-reports on team performance in reality. In future research, one could test different possibilities for detecting false self-reports.

Evidence from tournament experiments shows that team members invest in activities that harm other teammates in order to improve their own relative standing (Harbring and Irlenbusch 2011). Future research may also investigate the reporting behavior in a setting where teammates may not only self-report their own efforts but are also able to transfer (possibly false) information about their fellows’ efforts.

**References**


Abeler, J., Becker, A., Falk, A., 2014. Representative evidence on lying costs. Journal of Public Economics, 113, 96-104.

Alchian, A., & Demsetz, H. 1972. Production, information costs, and economic organization. The American Economic Review, 62(5), 777–795.

Boyle, M. 2001. Performance reviews: Perilous curves ahead. Fortune 143, 187-188.

Fischbacher, U. 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Experimental Economics, 10(2), 171–178.

Campbell, J. and Lee, C. 1988. Self-Appraisal in Performance Evaluation: Development versus Evaluation. The Academy of Management Review, 13(2), 302-314.

Chaudhuri, A. 2011. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. Experimental Economics, 14(1), 47–83.

Chen, Z., Kemp, S. 2012. Lie hard: The effect of self-assessments on academic promotion decisions. Journal of Economic Psychology, 33(3), 578–589.

Conrads, J., Irlenbusch, B., Rilke, R. M., Schielke, A., & Walkowitz, G. 2014. Honesty in tournaments. Economics Letters, 123(1), 90–93.

Dunning, D., Heath, C., & Suls, J. M. 2004. Flawed self-assessment: Implications for health, education, and the workplace. Psychological Science in the Public Interest, 5, 69–106.

Engel, C., Rockenbach, B. 2010. We Are Not Alone: The Impact of Externalities on Public Good Provision, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Voluntary Cooperation, No. E15-V3

Erat, S., Gneezy, U. 2012. White lies. Management Science, 58(4), 723-733.

Fischbacher, U., & Gächter, S. 2010. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. The American Economic Review, 100 (1), 541–556.

Fischbacher, U., Föllmi-Heusi, F., 2013. Lies in Disguise – An experimental study on cheating, Journal of the European Economic Association, 11(3), 525–547.

Gibson, R., Tanner, C., & Wagner, A. F. 2013. Preferences for Truthfulness: Heterogeneity among and within Individuals. The American Economic Review, 103(1), 532–548.

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

Gneezy, U., Rockenbach, B., & Serra-Garcia, M. 2013. Measuring lying aversion. Journal of Economic Behavior & Organization, 93(0), 293–300.

Greiner, B., 2004. An Online Recruiting System for Economic Experiments. In: Kurt Kremer, Volker Macho (eds.). Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63, Göttingen: Ges. Für Wiss. Datenverarbeitung, 79-93.

Harbring, C., Irlenbusch, B. 2011. Sabotage in Tournaments: Evidence from a Laboratory Experiment. Management Science 57, 611-627.

Hoffmann, M., Lauer, T., & Rockenbach, B. 2013. The royal lie. Journal of Economic Behavior & Organization, 93(0), 305–313.

Kaiser, J. 2007. An exact and a Monte Carlo proposal to the Fisher–Pitman permutation tests for paired replicates and for independent samples. The Stata Journal, 7(3), 402-412.

Lauer, T. 2011. Feedback and leading by example [Working Paper]. Faculty of Economics, Law and Social Sciences, University of Erfurt.

Milkovich, G. T., and Newman, J. M. 2008. Compensation ( 9th editon ). New York : McGraw - Hill/Irwin

Prendergast, C. 1999. The provision of incentives in firms. Journal of Economic Literature, 7-63.

Shore, T. H., & Tashchian, A. 2002. Accountability forces in performance appraisal: Effects of self-appraisal information, normative information, and task performance. Journal of Business and Psychology, 17, 261–274.

Vanberg, C. 2008. Why do people keep their promises? An experimental test of two explanations, Econometrica, 76, 1467-1480.

## Appendix

### Instructions for the experiment (for the treatment ENDOGEN-SR)

---

### General information

We welcome you to this economic experiment. It is very important that you read the following explanations carefully. If you have any questions, please address them to us. In this experiment, you can earn money depending on your own and other participants' decisions.

During the experiment, you are not allowed to talk to other participants. Non-compliance with this rule will result in your being excluded from the experiment and from all payments. All decisions will be taken anonymously, i.e., none of the other participants will learn the identity of the participant who has taken a specific decision. Payment is anonymous as well, i.e., no participant knows the payments of other participants.

During the experiment, your entire income is calculated in points. At the end of the experiment, the total number of points earned during the experiment are converted to Euros, where

35 points = 1 Euro.

At the end of today's experiment, you will receive the number of points earned during the experiment plus a 2,50 € show-up fee.

On the following pages, the detailed procedure of the experiment is explained.

### Course of the experiment

- The experiment will consist of **20 rounds**, each round having the same structure.
- You will be part of a group of 4 members. During the whole experiment, you will be exclusively interacting with the members of your group. The composition of the group will remain unchanged across all rounds of the experiment.

### Contributors and observer

Your group will consist of **three contributors** and **one observer**. Whether you will be assigned to the role of one of the contributors or the observer will be randomly determined and communicated to you before the first round begins. Moreover, a letter (A, B, or C) will be randomly assigned to each contributor. This letter will be visible to the other contributors and to the observer, and will remain unchanged throughout the whole experiment. At the beginning of each round, each **contributor** will receive an **endowment** of **20 points**. The **observer** will receive an **endowment** of **10 points**.

**Stage 1: Decisions of the contributors**

- Each **contributor** has to decide how many of the 20 points she wants to **contribute** to the **group account.** Points which are not contributed to the group account will be kept by the contributor. Possible contributions are: **[0, 5, 10, 15, or 20] points**.
- Each point contributed will be **multiplied by two,** and the **sum** of the doubled points is the **group result**. (sum of contributed points x 2 = group result)
- **All contributors** will be **informed** about the **individual decisions** and the **group result**.
- **The observer** will only be **informed** about the level of the **group result**.
- **Each contributor will decide** what information about the level of her contribution is **transmitted** to the **observer**.

**Distribution of the group result – see figure below**

- The group result is divided into three parts:
  - **Part I,** ¼ of the group result, is distributed equally among the **3 contributors**.
  - The **distribution** of **part II**, ½ of the group result, is carried out by the **observer**.
  - **Part III**, ¼ of the group result, is **given to the observer**.

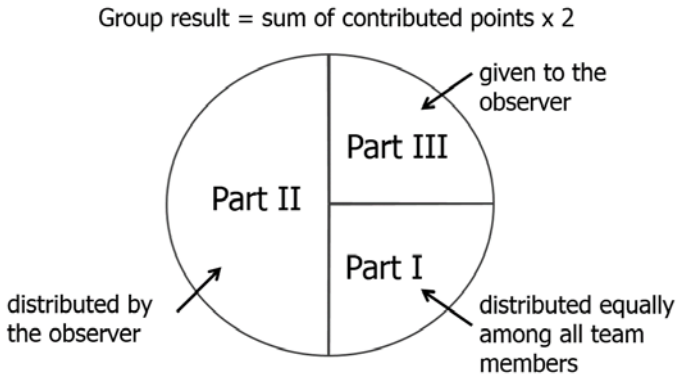



Figure: Distribution of the group result

**Stage 2: Decisions of the observer**

- The observer **cannot contribute** her round endowment to the group account.
- The observer has to **completely** distribute **part II among the contributors**. The observer cannot keep any points for herself.
- **At her own discretion**, the observer assigns a share of part II to each contributor.

**Calculation of the round incomes of the contributors**

- The round income of each contributor consists of three parts:
  - **Share of the round endowment** which has not been contributed,
  - **Share of part I** which has been distributed equally among all contributors,
  - **Share of part II** which the observer has assigned to the contributor.

**Round income of the contributor =**

+ 20 – contribution + part I / 3 + assigned share of part II

**Calculation of the round income of the observer**

- The round income of the observer consists of two parts:
  - **Round endowment** and
  - **part III**

**Round income of the observer =** 10 + part III

**Stage 3: End of the round**

- At the end of each round, all contributors will be informed about the decision of the observer and about their round incomes.
- The observer will be informed about her round income.

**Total income**

Your total income will be the sum of the incomes from all rounds.