Do different people report the same social norms?

Geoffrey Castillo[*]   Lawrence Choo[†]   Veronika Grimm[‡][§]

6th July 2022

If the Krupka-Weber (2013) norm-elicitation task captures pre-existing social norms, then the elicited norms should be independent of one's role in a game or one's social preferences. We test this idea in a complex game that features rich interactions. We find that different people, even when they have conflicting incentives, report the same social norms. Our results further validate the use of the Krupka-Weber task to measure social norms.

Keywords: social norms, norm elicitation, laboratory experiment, methodology, ultimatum game
JEL codes: C91, C92, C71, D90

1 Introduction

The Krupka-Weber (2013) norm-elicitation task is a popular tool to measure social norms in experiments.[1] In this task, subjects rate the social appropriateness of various actions in a game. Subjects earn a bonus if they give the same response as most other subjects. The norm is thus the focal point of a coordination game.

Given its popularity, many have worked to assess the robustness of this task. For example, in the usual Krupka-Weber norm-elicitation experiment, different subjects play the game and report the corresponding norms norms: Erkut et al. (2015) found that subjects who play the game report the same norms as those who do not; and D'Adda et al. (2016), that playing the game or reporting the norms first leads to similar reports. Another concern is that the task measures focal points unrelated to norms: Fallucchi and Nosenzo (2022) showed that competing focal points do not change the reports.

We continue to stress-test the Krupka-Weber norm-elicitation task and push it further. If the task really captures social norms—a construct that pre-exists the interactions in the game—then different people should report the same norms, even if their incentives or their preferences tell them not to. The robber and their victim should both say that stealing is socially inappropriate, even if downplaying stealing would make the robber appear like a better person. To test this idea, we set up a complex game and compare the norms reported by people who have different stakes in the game, a different understanding of the game, and different social preferences. We find that the reported norms do not differ significantly on these points, providing further validation for the use of the Krupka-Weber norm-elicitation task.

We rely on the whistleblowing game of Choo et al. (2019), an ultimatum game with asymmetric information played between two teams of three players. A pot of money contains one of three possible amounts; only Team A observes the exact amount. In Team A the Proposer makes a proposal on how to split the money between the two teams. The remaining two members of Team A can whistleblow to inform Team B of the amount in the pot. We compare two versions of the game by manipulating whether one or two members of Team A can whistleblow.

We look at the social norm behind the proposal and the one behind the whistleblowin decision. All subjects, irrespective of their role, report all norms. While subjects experience only one version of the game—one vs two whistleblowers—we also ask them to submit norms for the version of the game they have not actually experienced.

We find that a subject's experience does not matter: subjects playing either version of the game report the same norms. We also find that a subject's role does not matter: Proposers, members of Team A and members of Team B all agree on

the social norms associated with the Proposer's proposal and with the decision to not whistleblow. We further classify subjects depending on how their decision compares with the average decision of all relevant subjects. We also classify them depending on their reported belief and their degree of altruism as revealed by the Social Value Orientation measure (Murphy et al., 2011). Again, none of these matter: subjects of different types perceive the same norms.

2 Experimental design

2.1 The game

We refer to Choo et al. (2019) for the full details on the game.

There are two teams of three players and a pot of money that contains 30€, 90€, or 150€ with equal probability. Team A knows exactly how much there is in the pot; Team B only knows that it can be one of these three amounts.

In Team A there is one Proposer and two Insiders, and in Team B there are three Outsiders. The Proposer proposes how to split the money between the two teams. Team A sees how much is kept for Team A and how much is offered to Team B while Team B only sees how much they are offered. Note that the Proposer can sometimes exploit the asymmetric information about the pot to unfairly benefit Team A at the expense of Team B (see Choo et al., 2019, for details).

Insiders and Outsiders then vote on whether to accept the proposal. It is accepted if the votes reach majority, otherwise all players get 0€. The money is shared equally between all members of a team.

After players have cast their vote but before the majority decision is implemented, the Insiders can whistleblow to reveal how much there is in the pot. In the *One Whistleblower (OneWB)* treatment, only one randomly chosen Insider can whistleblow to reveal the amount in the pot. On the other hand, in the *Two Whistleblowers (TwoWB)* treatment, both Insiders can whistleblow and Team B is informed it if at least one Insider does so. In both treatments whistleblowing costs the Insider 0.5€.

We use the strategy method to elicit the initial decisions of the Proposer and of the Insiders. Outsiders, since they do not know the amount in the pot, submit only one acceptance strategy.

3

2.2 The norms

To measure the social norms we use the norm-elicitation task of Krupka and Weber (2013). We look at the norm underlying the Proposer's proposal and, conditional on the Proposer's proposal, at the norm underlying the Insider's decision to not whistleblow. For the latter, in the TwoWB treatment we control for subjects' beliefs by asking them to report the norm in two situations: when they believe the other is going to whistleblow and when they believe the other is not going to whistleblow. So in total we ask for four sorts of norms.

We concentrate on a few proposals: 0€, 12€, 15€, 36€, 45€, 60€, and 75€. These proposals amount to 0%, 40% and 50% of the pots, which are salient numbers in ultimatum and dictator games (Camerer, 2003). We use all these proposals for all pots, whenever possible. As a result, we use 17 pot-proposal combinations, which means that, since we have four sorts of norms, we ask for 68 norms in total (see Appendix C.1 for a list).

All subjects report all norms irrespective of their role and of the treatment they have been assigned to. We can thus make two comparisons:

Experience: Do subjects experiencing different treatment conditions report the same norms about a given treatment?

Role: Do subjects playing different roles in the game report the same norms?

We also ask subjects to report their beliefs regarding an Insider's decision to whistleblow and elicit their social value orientation (Murphy et al., 2011).[2] We use this information to classify subjects into different types and see if different types of subjects report similar norms:
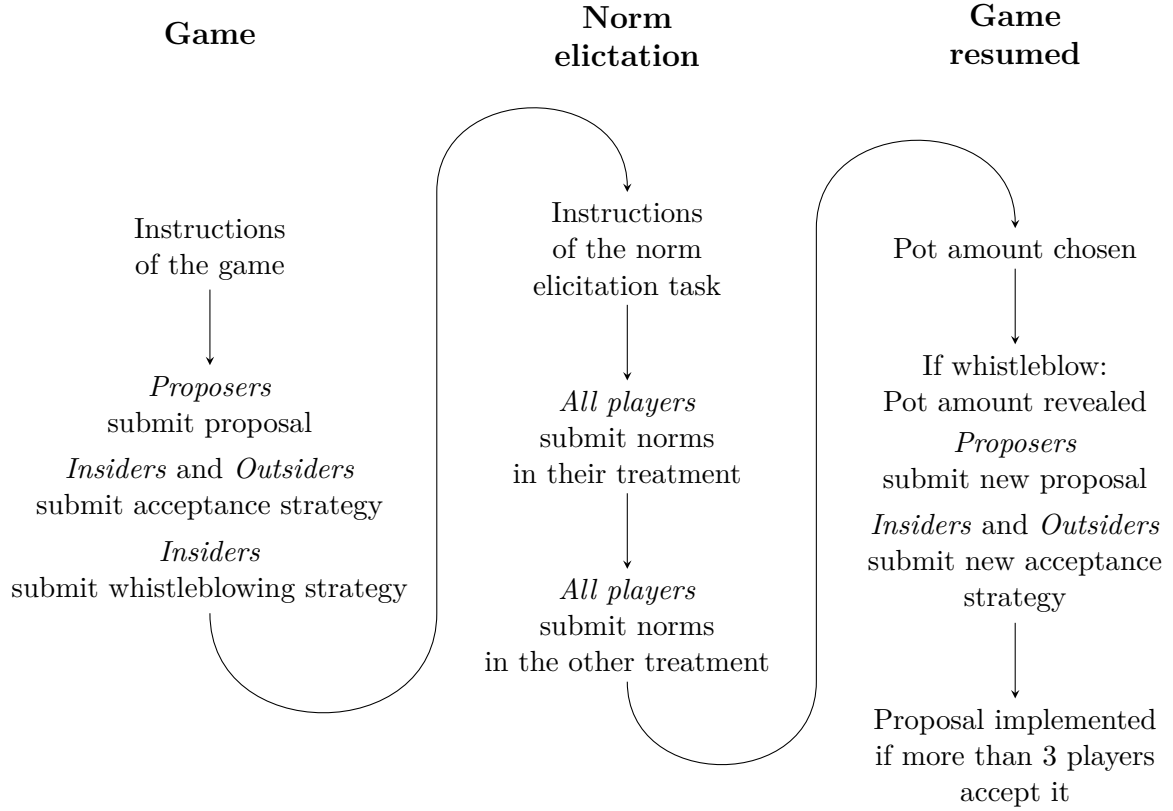
Type: Do different types of subject report the same norms?

2.3 Procedure

The procedure of the experiment is summarised in Figure 1. At the beginning of the experiment subjects were told that the experiments had two parts but they only had instructions for the first one—the game. The instructions featured

---

[2]We used the zTree code from Crosetto et al. (2012).

Figure 1: Summary of the experiment.

control questions. Then, the game started and subjects submitted their decisions. ⁹⁹
Subjects were also asked to report their beliefs about the whistleblowing strategy ¹⁰⁰
of a Member of Team A. This belief elicitation was not incentivised. ¹⁰¹

Before their decisions were implemented, the game was stopped and the second ¹⁰²
part started. The norm-elicitation task was introduced, a new set of instructions ¹⁰³
was distributed and subjects reported the norms about the game in their own ¹⁰⁴
treatment. They were explained the difference between their treatment and the ¹⁰⁵
other treatment, and they reported the norms about the norms in the other ¹⁰⁶
treatment. Once this was over, the game resumed: subjects' decisions from the ¹⁰⁷
first part were implemented and the subjects were told the outcome of the game. ¹⁰⁸

In the game part subjects could earn their earnings from the game. In the norm ¹⁰⁹
elicitation part the computer randomly selected for all subjects one of the norms ¹¹⁰
and looked for the social appropriateness rating most often reported. Subjects ¹¹¹

Figure 1: Summary of the experiment.

control questions. Then, the game started and subjects submitted their decisions. [99]
Subjects were also asked to report their beliefs about the whistleblowing strategy [100]
of a Member of Team A. This belief elicitation was not incentivised. [101]

Before their decisions were implemented, the game was stopped and the second [102]
part started. The norm-elicitation task was introduced, a new set of instructions [103]
was distributed and subjects reported the norms about the game in their own [104]
treatment. They were explained the difference between their treatment and the [105]
other treatment, and they reported the norms about the norms in the other [106]
treatment. Once this was over, the game resumed: subjects' decisions from the [107]
first part were implemented and the subjects were told the outcome of the game. [108]

In the game part subjects could earn their earnings from the game. In the norm [109]
elicitation part the computer randomly selected for all subjects one of the norms [110]
and looked for the social appropriateness rating most often reported. Subjects [111]

could earn 20€ if they had reported the same rating. Finally, for each subject the computer randomly selected one of the two parts and subjects were paid their earning in this part.

Before exiting the laboratory subjects completed the social value orientation measure and answered standard demographic questions. Appendix C.2 gives more details on how the experiment was implemented and Appendix C.3 reproduces the instructions.

2.4 Implementation

We conducted the experiment between July and November 2017 at the LERN laboratory (FAU, Nürnberg). We randomly recruited with ORSEE (Greiner, 2015) 126 subjects, 63 in each treatment, and have data for 21 Proposers, 42 Insider, and 57 Outsiders.[3] We programmed the experiment using zTree (Fischbacher, 2007). A session lasted about 1 hour and 30 minutes for an average payment—including the show-up fee of 4€—of 15.5€ (SD $\simeq$ 10€).

3 Results

We performed a power analysis of our statistical tests given our sample size. With a significance level $\alpha = 0.10$ and a power $1 - \beta = 0.80$ the analysis reveals that we are able to detect medium to large effect (effect size d between 0.6 and 0.7, effect size f between 0.2 and 0.3; Cohen, 1988). Appendix A details our power analysis.

Below we only report non-parametric tests. Appendix B.3 uses econometrics and finds the same results.

3.1 Norms are independent of experience

We first look at whether subjects in different treatments report the same norm about a given treatment. Table 1 shows the results of this comparison. For example, in the top-left part of the Table we look at the social appropriateness of not whistleblowing in the OneWB treatment—when only one Insider can whistleblow—and we compare

---

[3]We should have 63 Outsiders but due to a programming error we did not record the decisions of 6 Outsiders, 3 in each Treatment.

the norm reported by Proposers who actually experienced the OneWB treatment to the norm reported by Proposers who experienced the TwoWB treatment. As the Table shows some differences are significant, but few of them, and all but one disappear when we control for repeated testing (using the procedure from Benjamini and Hochberg, 1995).[4]

Given this result, we will pool the data generated by the two treatments to increase power for the next comparisons.

3.2 Norms are independent of role

We then look at whether subjects playing different roles in the game report the same norms. Figure 2 compares the norms reported by Proposers, Insiders, and Outsiders. The first column of the graph looks at the social appropriateness of the Proposer's proposal. The last three columns look at the social appropriateness of not whistleblowing in the OneWB treatment, in the TwoWB treatment when the other whistleblows, and in the TwoWB treatment when the other does not whistleblow. We can see that on average all three roles reported the same norm: Out of 54 comparisons, 4 are significantly different at the 10% level, 4 at the 5% level, and 1 at the 1% level. Significant differences, if any, seem to happen more at the extremes. Even then, we never observe the social norm jumping from, say, very socially inappropriate to very socially appropriate; and all but one of the significances disappear as soon as we apply the correction from Benjamini and Hochberg (1995). Table B1 in Appendix B.1 reports all p-values.

3.3 Norms are independent of type

Finally, we look at whether subjects of different types report the same norm. We have too few observations on the Proposers for results based on subtypes to be meaningful. Here we concentrate on the Outsiders—corresponding analyses for the Insiders give similar results and can be found in Appendix B.2. Table B2 in the same Appendix reports the full results.

We first look at the minimum amount offered to Team B under which they would reject the Proposer's proposal and do a mean split. Outsiders above the mean are

---

[4] We use the Stata command `multproc` from Newson and The ALSPAC StudyTeam (2003).

Table 1: p-values of two-sided Mann-Whitney tests comparing the social appropriateness of not whistleblowing between the OneWB and the TwoWB treatments.

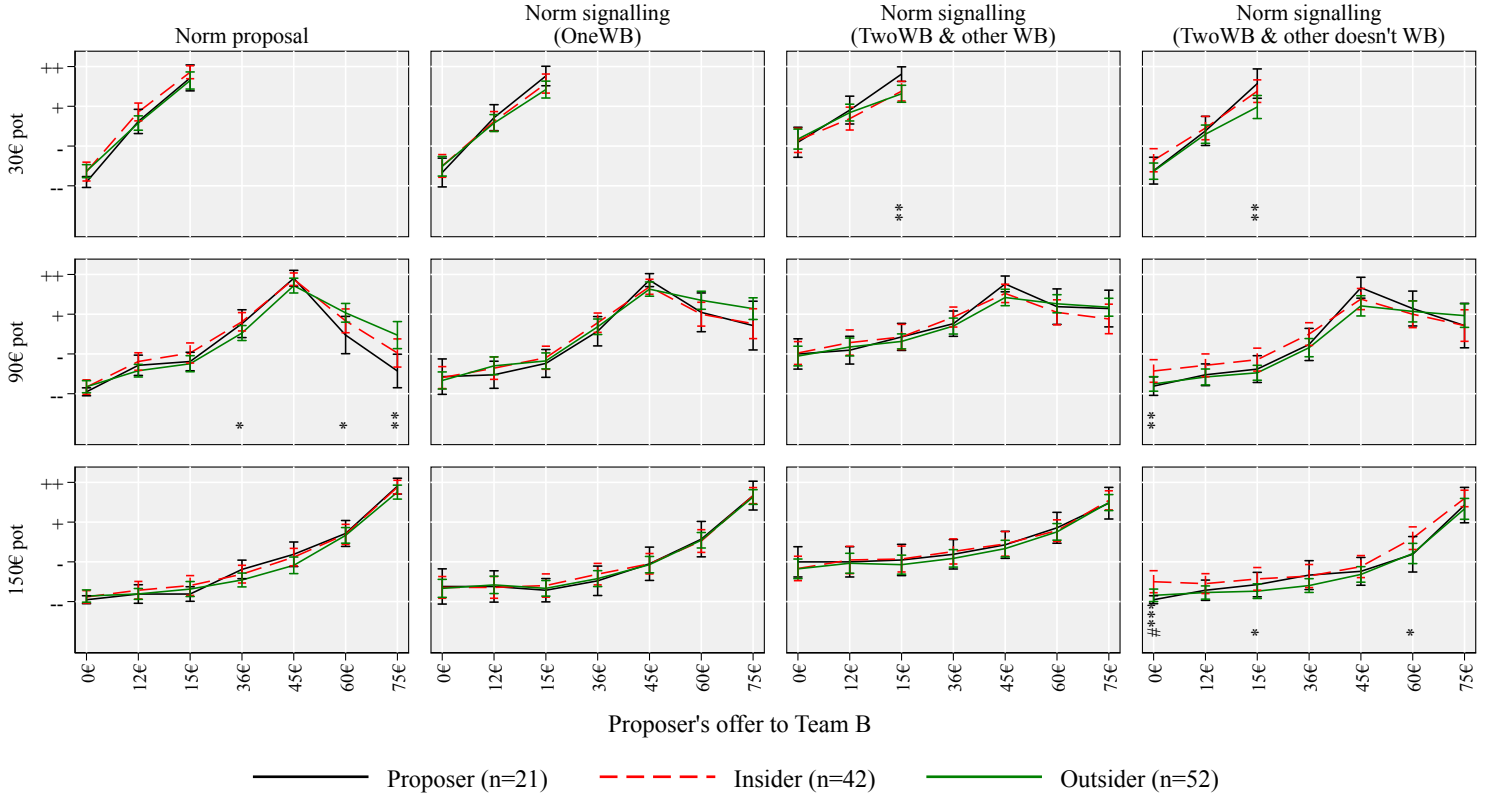| Proposal	Norm whistleblowing, OneWB treatment			Norm whistleblowing, TwoWB treatment, other whistleblows			Norm whistleblowing, TwoWB treatment, other does not whistleblow		
	30€	90€	150€	30€	90€	150€	30€	90€	150€
Norms submitted by Proposers ($n = 21$ in each cell)									
0	0.20	0.18	0.06*	0.88	0.85	0.85	0.37	0.06	0.29
12	0.91	0.50	0.24	0.60	0.84	0.85	0.79	0.97	0.81
15	0.04**	0.18	0.28	0.04**	0.48	0.68	0.51	0.76	0.80
36		0.22	0.22		0.51	0.91		0.45	0.85
45		0.08*	0.79		0.17	0.44		0.53	0.82
60		0.06*	0.12		0.16	0.52		0.01***	0.88
75		0.28	0.92		0.01***	0.67		0.14	0.64
Norms submitted by Insiders ($n = 42$ in each cell)									
0	0.70	0.18	0.04*	0.31	0.23	0.63	0.41	0.76	0.79
12	0.02**	0.08*	0.17	0.49	0.12	0.30	0.85	0.78	0.27
15	0.11	0.13	0.09*	0.61	0.45	0.76	0.40	0.35	0.29
36		0.24	0.22		0.65	0.43		0.40	0.58
45		0.22	0.30		0.42	0.23		0.36	0.96
60		0.26	0.40		0.75	0.77		0.61	0.42
75		0.71	0.74		0.67	0.60		0.43	0.38
Norms submitted by Outsiders ($n = 57$ in each cell)									
0	0.78	0.77	0.57	0.32	0.10	0.04**	0.93	0.19	0.74
12	0.50	0.08*	0.60	0.73	0.30	0.09*	0.30	0.55	0.60
15	0.11	0.07*	0.31	0.62	0.52	0.46	0.00***†	0.09*	0.87
36		0.38	0.02**		0.64	0.37		0.59	0.34
45		0.34	0.30		0.77	0.83		0.01**†	0.17
60		0.73	0.03**		0.57	0.99		0.16	0.22
75		0.88	0.04**		0.41	0.95		0.03**	0.07*

Notes. p-values of two-sided Mann-Whitney tests.
Uncorrected p-values: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Significant p-values at the 10% level using Benjamini and Hochberg (1995) correction (within norm) with $q = 0.15$: †
For example, the first number, 0.20, is the p-value of a test with H_0: the average social appropriateness of not whistleblowing in the OneWB treatment when the pot is 30€ and the Proposer proposes 0€ to Team B is the same between subjects in the OneWB treatment and subjects in the TwoWB treatment.
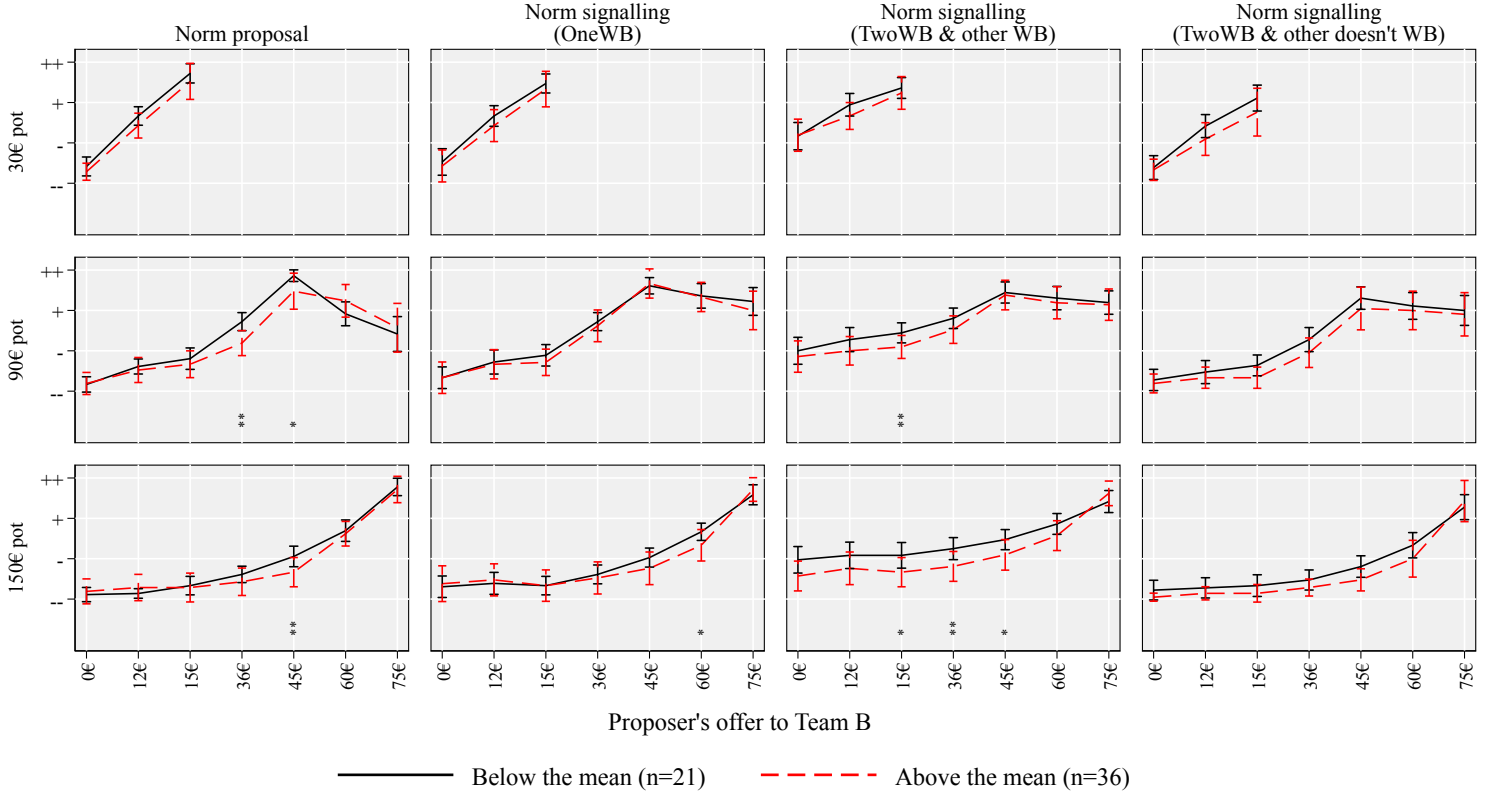
Notes. −−: very socially inappropriate; −: somewhat socially inappropriate; +: somewhat socially appropriate; ++: very socially appropriate.

∗ and # and correspond to Kruskal-Wallis tests, H_0: the average social appropriateness is the same across Proposers, Insiders, and Outsiders.

Uncorrected p-values: ∗ is $p < 0.01$, ∗∗ is $p < 0.05$, and ∗∗∗ is $p < 0.1$.

The only p-value significant after applying Benjamini and Hochberg (1995) correction with $q = 0.15$ is denoted by #.

Figure 2: Mean and 95% confidence interval of the social appropriateness of the Proposer's offers to Team B for different player roles.

subjects who expect a lot from the Proposer while subjects below are subjects who [167]
would be content with a smaller amount. This split thus separates Outsiders with [168]
different expectations. We then compare the norm reported by subjects above and [169]
below the mean. Figure 3 shows that these two groups of Insiders essentially report [170]
the same norms. [171]

We then use the social value orientation measure to classify the Outsiders into [172]
different types. In our sample, according to the classification of Murphy et al. [173]
(2011), about 65% of the Outsiders are 'prosocial' and 35% are 'individualist'. We [174]
compare the norms reported by the two types of Outsiders. Figure 4 shows that [175]
prosocial Outsiders and individualist Outsiders also report the same norms. [176]

References

Benjamini, Yoav and Yosef Hochberg (1995), "Controlling the false discovery rate: A practical and powerful approach to multiple testing." *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300.

Camerer, Colin F. (2003), *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton University Press, Princeton.

Choo, Lawrence, Veronika Grimm, Gergely Horváth, and Kohei Nitta (2019), "Whistleblowing and diffusion of responsibility: An experiment." *European Economic Review*, 119, 287–301.

Cohen, Jacob (1988), *Statistical Power Analysis for the Behavioural Sciences*, second edition. Lawrence Erlbaum Associates.

Crosetto, Paolo, Ori Weisel, and Fabian Winter (2012), "A flexible z-Tree implementation of the social value orientation slider measure." Jena economic research paper no. 2012 – 062, Friedrich Schiller University and Max Planck Institute of Economics, Jena.

D'Adda, Giovanna, Michalis Drouvelis, and Daniele Nosenzo (2016), "Norm elicitation in within-subject designs: Testing for order effects." *Journal of Behavioral and Experimental Economics*, 62, 1–7.

Erkut, Hande, Daniele Nosenzo, and Martin Sefton (2015), "Identifying social norms using coordination games: Spectators vs. stakeholders." *Economics Letters*, 130, 28–31.

Fallucchi, Francesco and Daniele Nosenzo (2022), "The coordinating power of social norms." *Experimental Economics*, 25, 1–25.

Figure 3: Mean and 95% confidence interval of the social appropriateness of the Proposer's offers to Team B for two Outsider types: Outsiders above and below the mean of the minimum amount offered to Team B under which the Proposer's proposal is rejected.

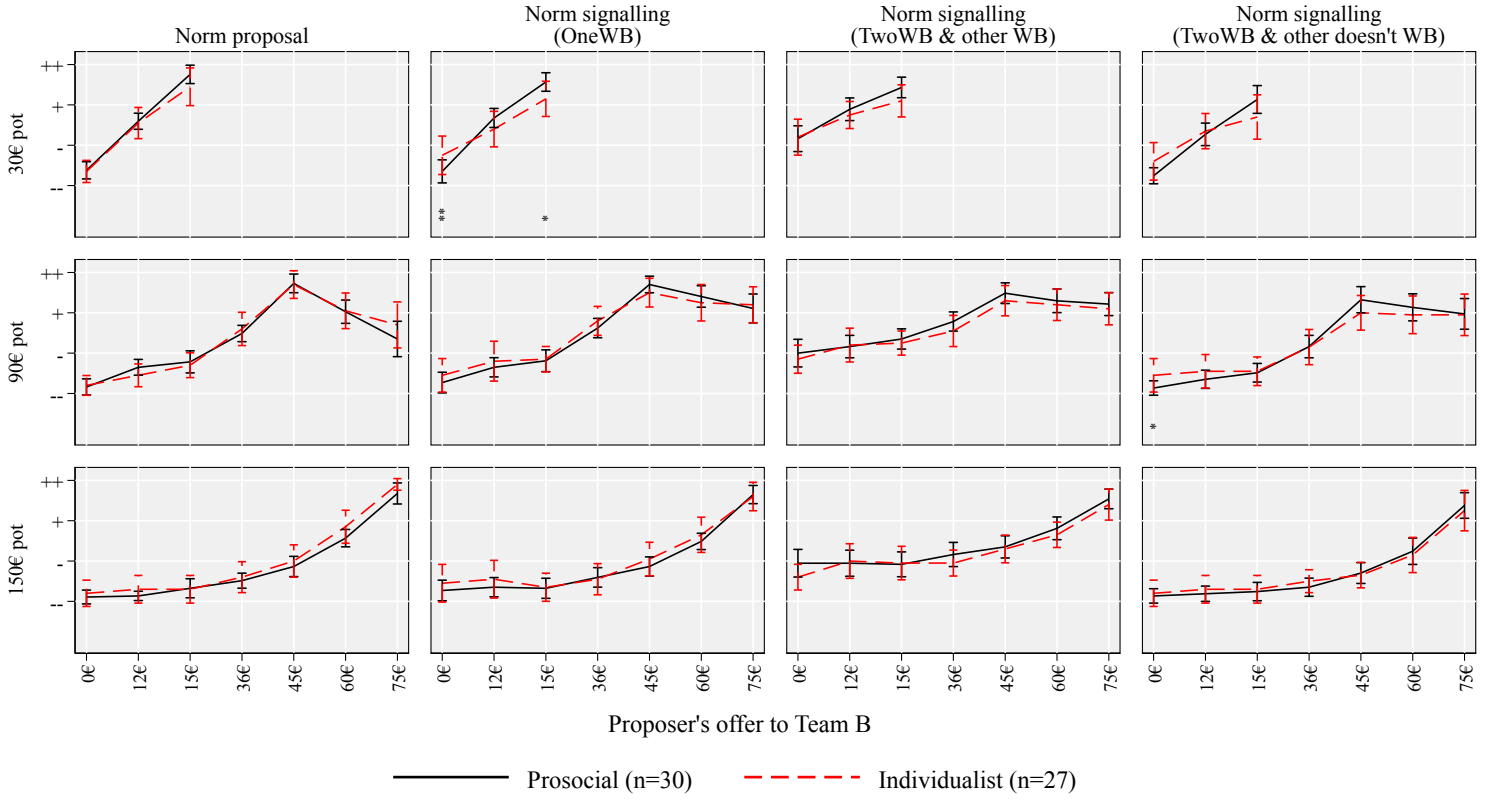*Notes.* 'Below the mean': weakly; 'Above the mean': strictly.
——: very socially inappropriate; —: somewhat socially inappropriate; +: somewhat socially appropriate; ++: very socially appropriate.
∗ corresponds to Kruskal-Wallis tests, H_0: the average social appropriateness is the same across Proposers, Insiders, and Outsiders.
Uncorrected p-values: ∗ is $p < 0.01$, ∗∗ is $p < 0.05$, and ∗∗∗ is $p < 0.1$.
None of the p-values are significant after applying Benjamini and Hochberg (1995) correction with $q = 0.15$.

11

Notes. −−: very socially inappropriate; −: somewhat socially inappropriate; +: somewhat socially appropriate; ++: very socially appropriate.

∗ corresponds to Kruskal-Wallis tests, H_0: the average social appropriateness is the same across Proposers, Insiders, and Outsiders. Uncorrected p-values: ∗ is $p < 0.01$, ∗∗ is $p < 0.05$, and ∗ ∗ ∗ is $p < 0.1$.

None of the p-values are significant after applying Benjamini and Hochberg (1995) correction with $q = 0.15$.

Figure 4: Mean and 95% confidence interval of the social appropriateness of the Proposer's offers to Team B for two Outsider types: Prosocial or Individualist according to the SVO measure.

Fischbacher, Urs (2007), "Z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics*, 10, 171–178.

Gächter, Simon, Leonie Gerhards, and Daniele Nosenzo (2017), "The importance of peers for compliance with norms of fair sharing." *European Economic Review*, 97, 72–86.

Gächter, Simon, Daniele Nosenzo, and Martin Sefton (2013), "Peer effects in pro-social behavior: Social norms or social preferences?" *Journal of the European Economic Association*, 11, 548–573.

Greiner, Ben (2015), "Subject pool recruitment procedures: Organizing experiments with ORSEE." *Journal of the Economic Science Association*, 1, 114–125.

Krupka, Erin L., Stephen Leider, and Ming Jiang (2017), "A meeting of the minds: Informal agreements and social norms." *Management Science*, 63, 1708–1729.

Krupka, Erin L. and Roberto A. Weber (2013), "Identifying social norms using coordination games: Why does dictator game sharing vary?" *Journal of the European Economic Association*, 11, 495–524.

Murphy, Ryan O., Kurt A. Ackermann, and Michel J. J. Handgraaf (2011), "Measuring social value orientation." *Judgment and Decision Making*, 6, 771–781.

Newson, Roger and The ALSPAC StudyTeam (2003), "Multiple–test procedures and smile plots." *Stata Journal*, 3, 109–132.

Appendices

Appendix A Power analysis

To perform a power analysis, assuming the two groups we compare are of equal sizes, $n_1 = n_2 = n$, we first use the equation from Kupper and Hafner (1989) and Lehr (1992):

$$n = 2\left(\frac{z_{1-\alpha/2} + z_{1-\beta}}{d}\right)^2 \tag{1}$$

where α is the probability of rejecting the null hypothesis when it is true (type I error), β the probability of not rejecting the null hypothesis when it is false (type II error), z_x the inverse normal value with $\Pr(z > z_x) = x$ when $z \sim \mathcal{N}(0,1)$, and d the effect size.

We use $\alpha = 0.1$ and $\beta = 0.2$, corresponding to a power of 0.8, giving us $z_{1-\alpha/2} = z_{0.05} = 1.645$ and $z_{1-\beta} = z_{0.8} = 0.841$. We cannot use previous studies to determine the effect size d so we rely on Cohen (1988) and select a 'medium' effect size $d = 0.5$. Plugging these numbers into the equation above and rounding to the nearest integer gives us $n = 49$. Lowering α to 0.05 increases n to 63.

We confirm these sample sizes using G*Power (Faul et al., 2007). For the power of our Mann-Whitney U tests we select 'Means: Wilcoxon-Mann-Whitney test (two groups)' and enter the same parameters as above. G*Power gives us $n = 53$ for $\alpha = 0.1$ and $n = 67$ for $\alpha = 0.05$, similar to what we got with the equation above.

We also use Kruskal-Wallis tests but computing the required sample size for the Kruskal-Wallis test is notoriously difficult, especially because we have no historical or pilot data (Fan and Zhang, 2012; Fan et al., 2011). The best we can do is to use the parametric equivalent of the Kruskal-Wallis test, the one-way ANOVA, which provides a lower-bound to the power of the Kruskal-Wallis test (Lachenbruch and Clements, 1991). Using the same parameters as above in G*Power and selecting a 'medium' effect size $f = 0.25$, also from Cohen (1988), gives us a total sample size of 126 with $\alpha = 0.1$ and 157 with $\alpha = 0.05$.

Table A1 summarises the required sample sizes for the different tests and for different effect sizes. The take-away is that, with the exception of the experience comparison for the Proposers, our experiment is generally able to detect medium to large effects ($d \geq 0.5$ and $f \geq 0.25$).

We can take the problem from the other end and look at the power we achieve given our sample size. We also rely on G*Power to perform this post-hoc analysis of power. Figure A1 reports the achieved power of our Mann-Whitney U tests reported in Table 1 given different effect sizes. Unsurprisingly with only 21 Proposers for a power of 80% we are able to detect only 'very large' effect sizes (Sawilowsky, 2009).

Table A1: Sample sizes for the tests used in the experiment.

| | Effect size d | | | | | | | | |
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Equation, $\alpha = 0.1$	1236	309	137	77	49	34	25	19	15
Equation, $\alpha = 0.05$	1569	392	174	98	63	44	32	25	19
G*Power, Mann-Whitney, $\alpha = 0.1$	1296	324	145	82	53	37	27	21	17
G*Power, Mann-Whitney, $\alpha = 0.05$	1645	412	184	104	67	47	35	27	21

| | Effect size f | | | | | | | | |
	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5
G*Power, Kruskal-Wallis, $\alpha = 0.1$	773	345	195	126	88	65	51	41	33
G*Power, Kruskal-Wallis, $\alpha = 0.05$	966	431	244	157	110	82	63	51	42

Notes. In the first panel, when the effect size is d, the Table reports the sample size of each group assuming equal groups, i.e. $n_1 = n_2 = n$. In the second panel, when the effect size is f, the table reports the total sample size.
Effect sizes from Cohen (1988): For d: $d = 0.2$: small; $d = 0.5$: medium; $d = 0.8$: large. For f: $f = 0.1$: small; $f = 0.25$: medium; $f = 0.4$: large.
'Equation' refers to equation (1) on p.14. For G*Power see Faul et al. (2007).

With Insiders and Outsiders we can detect medium to large effects.

Figure A2 then reports the achieved power of our Kruskal-Wallis tests used in Figure 2 in the main text and Table B1 in the Appendix. With 120 subjects in total and for a power of 80% we are able to detect medium to large effect sizes.

Appendix B Detailed results of the norms

B.1 Role

Table B1 reports the p-values corresponding to Figure 2 in the main text.

B.2 Type

Figures B1 and B2 reproduce the analysis of Outsiders types reported in Figures 3 and 4 in the main text for Insiders. Table B2 reports the corresponding z-scores and p-values, aggregated at the pot level.

B.3 Econometrics

We now turn to econometrics. Since the norms are ordered—'very socially inappropriate' is less socially appropriate than 'somewhat socially inappropriate'—we
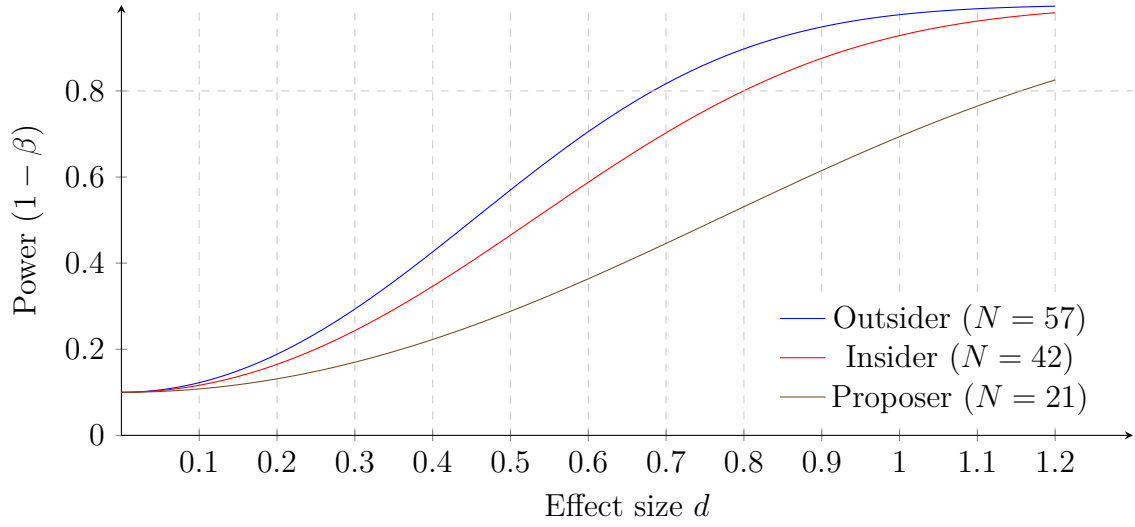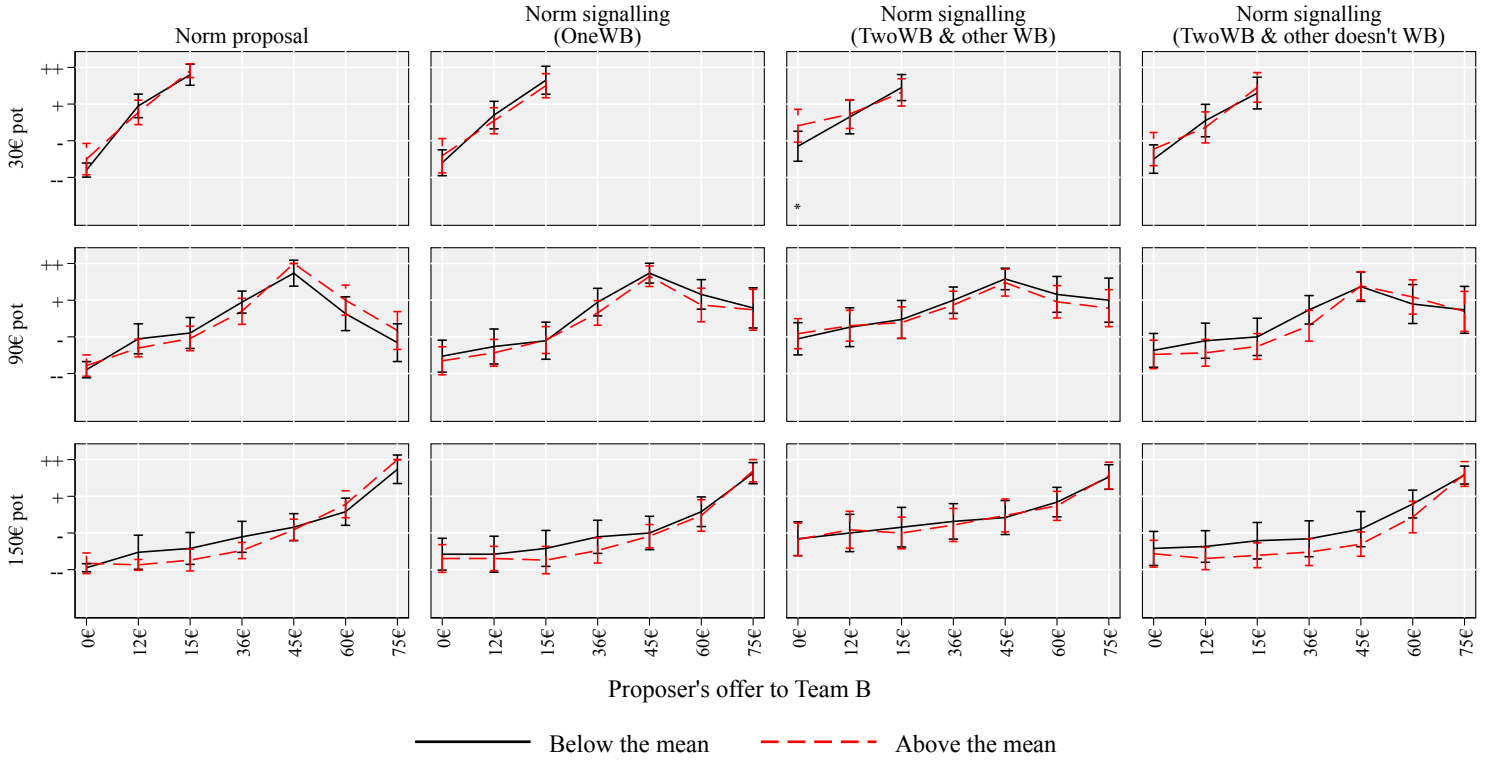
Figure A1: Power of the Mann-Whitney U tests used in Table 1 in the main text, $\alpha = 0.10$.
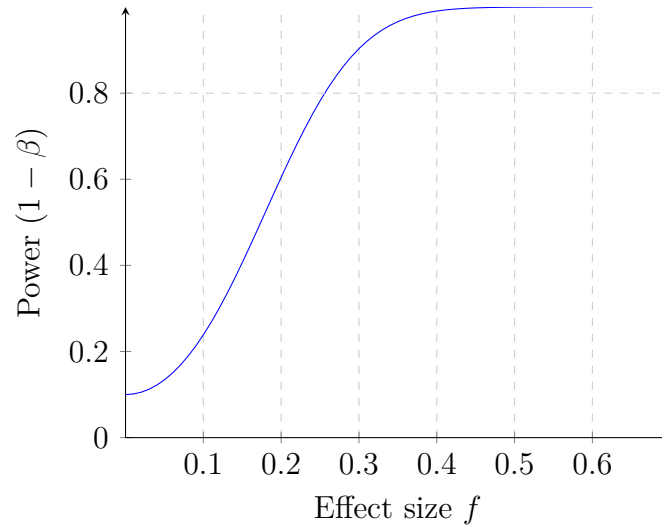
Figure A2: Power of the Kruskal-Wallis tests used in Figure 2 in the main text and Table B1 in the Appendix, $\alpha = 0.10$.

There were n=47 insiders. Amongst these, 20, 19 and 19 were below the mean when the pot contains 30€, 90€ and 150€, respectively.

Notes. ——: very socially inappropriate; −: somewhat socially inappropriate; +: somewhat socially appropriate; ++: very socially appropriate.

Letters correspond to Kruskal-Wallis tests, H_0: the average social appropriateness is the same across the two groups.

Uncorrected p-values: a is $p < 0.1$, b is $p < 0.05$, and c is $p < 0.01$.

None of the p-values are significant after applying Benjamini and Hochberg (1995) correction with $q = 0.15$.

Figure B1: Type comparison of social appropriateness, Insiders above and below the mean of the minimum amount offered to Team A under which the Proposer's proposal is rejected.
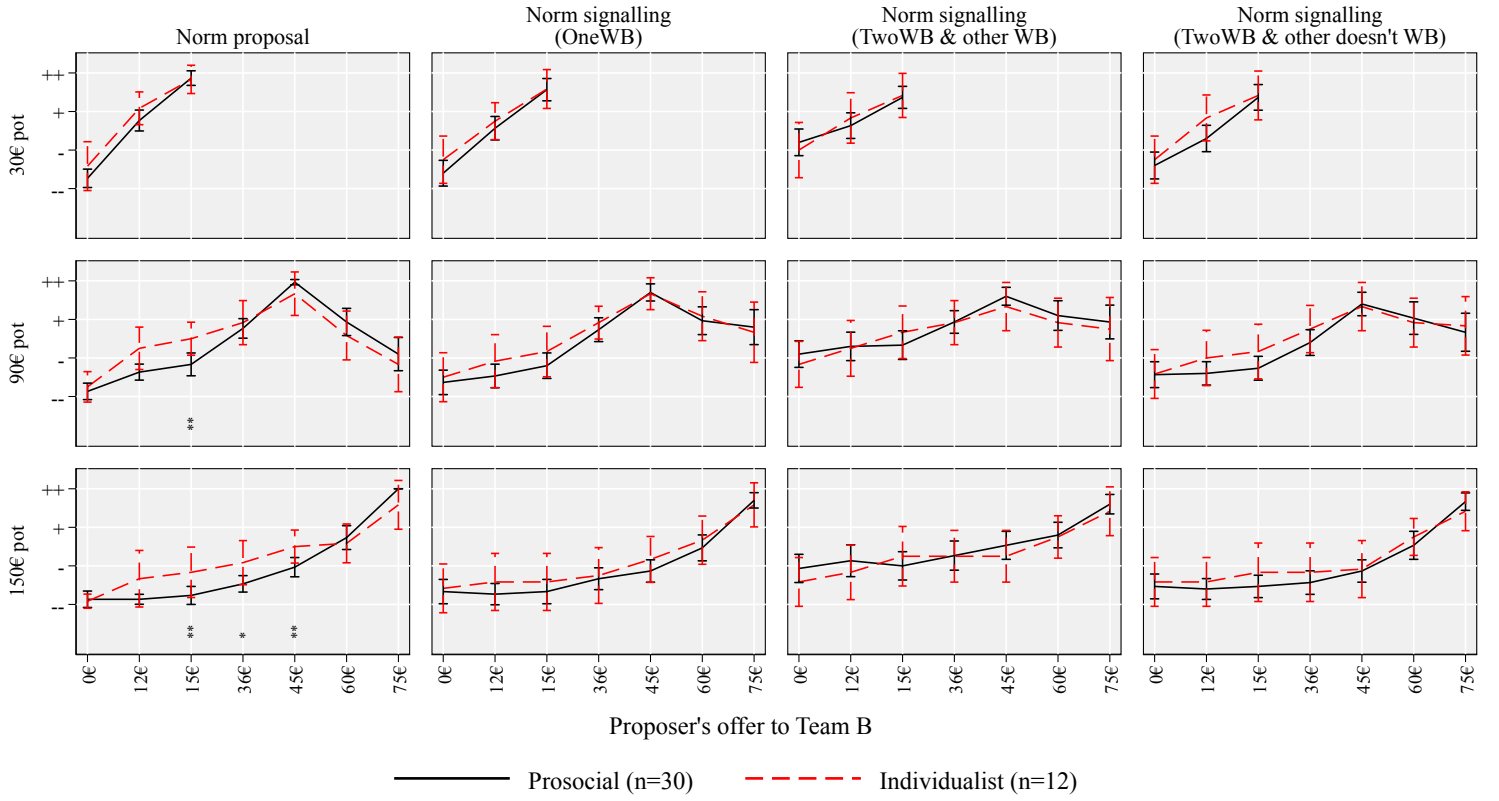
Figure B2: Type comparison of social appropriateness, Insiders categorised by the SVO measure as Prosocial or Individualist.

Notes. −−: very socially inappropriate; −: somewhat socially inappropriate; +: somewhat socially appropriate; ++: very socially appropriate.

Letters correspond to Kruskal-Wallis tests, H_0: the average social appropriateness is the same across the two groups.

Uncorrected p-values: a is $p < 0.1$, b is $p < 0.05$, and c is $p < 0.01$.

All p-values remain significant at the 5% level after applying Benjamini and Hochberg (1995) correction with $q = 0.15$.

Table B1: *p*-values corresponding to the role comparisons.

Proposal	Norm proposal			Norm whistleblowing, OneWB treatment			Norm whistleblowing, TwoWB treatment, Other whistleblows			Norm whistleblowing, TwoWB treatment, Other does not whistleblow		
	30€	90€	150€	30€	90€	150€	30€	90€	150€	30€	90€	150€
0	0.14	0.72	0.92	0.70	0.89	0.99	0.23	0.05^{**}	$0.01^{***}_{\dagger}$	0.96	0.88	0.53
12	0.19	0.24	0.85	0.67	0.48	0.62	0.59	0.17	0.20	0.70	0.78	0.95
15	0.23	0.21	0.70	0.14	0.92	0.98	0.03^{**}	0.19	0.07^{*}	0.03^{**}	0.76	0.76
36		0.08^{*}	0.12		0.40	0.65		0.15	0.33		0.30	0.72
45		0.17	0.15		0.41	0.99		0.16	0.61		0.23	0.81
60		0.08^{*}	0.92		0.16	0.99		0.92	0.10^{*}		0.62	0.82
75		0.01^{**}	0.16		0.24	0.88		0.61	0.56		0.61	0.87

Notes. *p*-values of Kruskal-Wallis tests.

Uncorrected *p*-values: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Significant *p*-values after applying Benjamini and Hochberg (1995) correction with $q = 0.15$: $\dagger$

For example, the first number, 0.14, is the *p*-value of a test with H_0: the average social appropriateness of a proposal of 0€ when the pot is 30€ is the same across Proposers, Insiders and Outsiders.

Table B2: *z*-scores corresponding to several type comparisons.

	Norm proposal			Norm whistleblowing, OneWB			Norm whistleblowing, TwoWB, Other whistleblows			Norm whistleblowing, TwoWB, Other does not whistleblow		
	30€	90€	150€	30€	90€	150€	30€	90€	150€	30€	90€	150€
Social value orientation:[a]												
Proposers	0.358	−0.942	−1.108	−0.083	−1.057	−0.944	−0.161	−1.214	0.196	1.756	0.039	1.155
Insiders	−1.282	−0.491	−1.860*	−0.626	−0.545	0.169	−0.941	−0.391	0.000	−0.155	0.112	0.628
Outsiders	0.723	−0.320	−0.676	0.703	−0.439	−0.093	0.434	−0.034	−0.430	0.881	0.688	0.361
All	−0.036	−0.991	−1.797*	0.032	−1.111	−0.486	−0.432	−0.623	−0.028	1.186	0.736	1.095
Mean splits: Minimum offered to Team B[b]												
Insiders[c]	−0.040	−0.344	0.704	0.797	0.888	0.433	−0.751	0.431	0.152	−0.368	0.507	1.527
Outsiders[d]	1.465	0.808	1.529	0.908	0.542	1.220	0.880	1.279	1.743	1.347	1.131	1.135

Notes. Two-sided Mann-Whitney U tests, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

[a] Prosocials vs. individualists, as defined in Murphy et al. (2011).
 A positive z-score means that on average prosocials reported higher social appropriateness scores.
 Number prosocials/individualists for Proposers: 15/6; Insiders: 30/12; Outsiders: 37/20; All: 82/38.
[b] A negative z-score means that on average subjects above the mean reported higher social appropriateness scores.
[c] In total: 42 Insiders. Number of Insiders above the mean: 22 in 30€ pot, 23 in 90€ pot, 23 in 150€ pot.
[d] In total: 57 Outsiders. Number of Outsiders above the mean: 21 for all pots.

Table B3: Results from the ordered logit model.

	Norm proposal			Norm whistleblowing, OneWB treatment			Norm whistleblowing, TwoWB treatment, Other does not whistleblow			Norm whistleblowing, TwoWB treatment, Other whistleblows		
	30€	90€	150€	30€	90€	150€	30€	90€	150€	30€	90€	150€
Experience[a]												
TwoWB	−0.02	0.03	0.13	−0.54**	−0.25	−0.60**	−0.41	−0.30	−0.14	−0.01	0.21	0.46
	(0.25)	(0.16)	(0.23)	(0.25)	(0.20)	(0.26)	(0.29)	(0.23)	(0.28)	(0.28)	(0.25)	(0.29)
Role[b]												
Outsider	−0.16	−0.10	−0.43	−0.20	0.12	−0.01	−0.53	−0.41	−0.83***	0.26	−0.05	−0.17
	(0.28)	(0.15)	(0.31)	(0.30)	(0.21)	(0.33)	(0.34)	(0.25)	(0.31)	(0.34)	(0.29)	(0.32)
Type												
SVO angle	−0.01	−0.01	−0.03***	−0.01	−0.01	−0.02	−0.02	−0.01	−0.02	−0.01	−0.01	−0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Acceptance strategy	−0.01	0.00	0.00	0.00	0.00	0.00	−0.01	0.00	−0.01	−0.01	−0.01	−0.01
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Belief whistleblow strategy	−0.02	0.00	0.00	−0.02	−0.01	0.00	−0.01	0.00	0.00	0.01	−0.01	−0.01
	(0.02)	(0.00)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
Proposer's proposal	0.41***	0.04***	0.10***	0.29***	0.05***	0.07***	0.21***	0.05***	0.06***	0.15***	0.04***	0.04***
	(0.06)	(0.00)	(0.01)	(0.03)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.00)	(0.00)
Observations	297.00	693.00	693.00	297.00	693.00	693.00	297.00	693.00	693.00	297.00	693.00	693.00
Pseudo R^2	0.34	0.09	0.35	0.23	0.15	0.25	0.17	0.14	0.21	0.09	0.08	0.11
Log-likelihood	−266.00	−856.00	−558.00	−314.00	−807.00	−653.00	−340.00	−812.00	−657.00	−359.00	−855.00	−842.00
Wald χ^2	63.70	101.37	118.39	82.64	128.06	91.64	92.55	108.85	108.99	77.99	86.77	94.14
Prob $> \chi^2$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Notes. Coefficients, standard errors in parentheses.
Standard errors clustered on subjects (99 clusters).
[*] $p < 0.1$, [**] $p < 0.05$, [***] $p < 0.01$.
[a]Base: subjects who participated in the OneWB treatment.
[b]Base: subjects who played as Insiders.

rely on an ordered logit model to analyse the norms reported by our subjects. We look at several comparisons at the same time: We compare subjects who played the TwoWB treatment to those who played the OneWB treatment; Insiders and Outsiders; and subjects who differ in their SVO angle, in the reported minimum amount of money offered to their team under which they would reject the Proposer's proposal, and in their belief about the whistleblowing strategy of an Insider.

The results of this regression are presented in Table B3. We verify our previous conclusions: few of these comparisons yield significant differences.

Appendix C Details on the experiment

C.1 List of norm questions

We look at 17 pot-proposal combinations:
- For the 30€ pot, proposals of 0€, 12€, 15€ = 3 combinations;
- For the 90€ and 150€ pots, proposals of 0€, 12€, 15€, 36€, 45€, 60€, 75€ = 7 combinations per for each pot.

For each pot-proposal combination, we study 4 kinds of norms:
- Norm behind the proposal by the Proposer,
- Norm behind the decision not to whistleblow in the OneWB treatment,
- Norm behind the decision not to whistleblow in the TwoWB treatment when the other whistleblows,
- Norm behind the decision not to whistleblow in the TwoWB treatment when the other does not whistleblow.

Therefore, in total we have 4 kinds of norms × 17 pot-proposal combinations = 68 norms.

C.2 Procedure

Subjects entered the laboratory and were randomly allocated a seat. The experimenter then started reading the instructions. The instructions first described the game in general terms and subjects answered questions controlling for their general understanding of the game. Then the instructions described how each role in the game (Proposer, Insider, Outsider) would make their decision. Once this was done groups were formed and subjects were informed of their role. They answered a second set of control question tailored to their particular role.

Then, it looked at the Proposer's proposal for this amount and at whether or not the Insiders chose to whistleblow. If they did not, the computer then looked at the voting strategy of the Insiders and the Outsiders. If they did, the precise amount in the pot was revealed to all.

C.3 Instructions

The next pages reproduces the instructions as they were seen by the subjects—they are here reproduced two-pages-on-one to save space. These instructions are for the OneWB treatment.

Instructions

Welcome to the experiment! Please switch off your electronic devices and remain silent. Also, do not communicate with the other participants. If you have a question at any time, raise your hand and an experimenter will come to your desk to answer it.

For your participation you will receive a fixed payment of 4€, and you may earn more depending on your choices and the choices of others. It is therefore crucial that you pay attention while we are reading these instructions.

The experiment consists of two tasks: the First Task and the Second Task. For each Task you will make one or more decisions. Decisions in one Task cannot affect the other Task. Importantly, only one of the two Tasks will ultimately determine your final payment. This will be done as follows: At the end of the experiment, the computer will randomly choose one of the Tasks for payment. If the First Task is selected, then we will pay you your earnings from the First Task on top of your fixed payment of 4€. If the Second Task is selected, we will pay you your earnings from the Second Task on top of the 4€. So, in each Task you should behave as if this was the one determining your final payment. In any case, the payment will be in cash and in private.

You will find below detailed instructions about the First Task. You will receive detailed instructions about the Second Task later, as we will explain soon.

First Task

1 Introduction

During the experiment you will interact with five other participants selected at random in this room. Then, you will be randomly placed in one of two teams, Team-A and Team-B. Each team will consist of three participants.

Participants in each team will be given a Role (Proposer of Team-A, Member of Team-A or Member of Team-B) which determines the type of decisions s/he will be making in this experiment. Here is how the roles are distributed:

Team-A = 1 x Proposer of Team-A + 2 x Members of Team-A

Team-B = 3 x Members of Team-B.

The next Section (Section 2) gives an overview of the experiment. Section 3 details how each Role makes their decisions.

2 Overview of the experiment

There is an envelope that can contain 30€, 90€ or 150€. Each amount is equally likely, meaning that the envelope can be 30€, 90€ or 150€ with equal chance. Once the amount of money in the envelope has been determined, the Participants in Team-A (one Proposer and two Members of Team-A) are informed of the amount of money in the envelope. The Members of Team-B are NOT informed of the amount of money in the envelope.

The experiment then proceeds as follows.

2.1. The Proposer plans how the money is to be distributed

The Proposer's task is to plan how to distribute the amount of money in the envelope between Team-A and Team-B.

The Proposer chooses the amount of money to be kept for Team-A, denoted by A€, and the amount of money to be given to Team-B, denoted by B€. Of course, A€ and B€ must add up to the amount present in the envelope. By choosing A€ and B€, the Proposer proposes that each participant in Team-A gets A€/3 and that each participant in Team-B gets B€/3.

2.2 The Members of Team A decide to send a "Signal" or not

After the Proposer has decided on A€ and B€:

- Members of Team-A are informed of both the amount kept for Team-A and the amount given to Team-B (they see A€ and B€)
- Members of Team-B only see the amount given to Team-B (they only see B€).

Then, ONE randomly chosen Member of Team-A will have the opportunity to inform the Members of Team-B of the amount of money in the envelope. S/he does so by deciding to send a "Signal". Sending the Signal is *anonymous* (e.g., the identity of the signal sender is never revealed) and cost the sender 0,50€.

The rest of the experiment depends on whether or not a Signal is sent.

2.2.1 No Signal is sent

If no Signal is sent by the randomly chosen Member of Team-A, each Member of Team-A must now decide if s/he would accept or reject the amount of money kept for Team-A (A€). Similarly, each Member of Team-B must decide if s/he would accept or reject the amount of money offered to Team-B (B€).

This implies that 5 participants must decide whether they accept or reject the distribution plan of the Proposer.

- If 3 or more participants accept the Proposer's distribution, A€ and B€ are implemented. Each participant in Team-A (including the Proposer) gets A€/3 and each participant in Team-B gets B€/3.
- If less than 3 participants accept the Proposer's distribution, A€ and B€ are NOT implemented. Each participant gets 0€.

2.2.2 A Signal is sent by the random Member of Team-A

If the random Member of Team-A sends a signal, the Members of Team-B are informed of the amount of money in the envelope.

The Proposer has then the opportunity to revise the amount proposed to Team-A and to Team-B. S/he decides on A€* (new amount for Team-A) and B€* (new amount for Team-B). By choosing A€* and B€*, the Proposer proposes that each participant in Team-A gets A€*/3 and that each participant in Team-B gets B€*/3. After the Proposer has decided on A€* and B€*:

- Members of Team-A and Members of Team-B see both A€* and B€*.

The Members of Team-A must now decide on whether they accept or reject the new amount of money the Proposer has proposed for Team-A (A€*). Similarly, the Members of Team-B must decide on whether they accept or reject the new amount of money offered to Team-B (B€*). This implies that a total of 5 participants accept or reject the new distribution plan of the Proposer:

- If 3 or more participants accept the Proposer's new distribution, A€* and B€* are implemented. Each participant in Team-A (including the Proposer) gets A€*/3 and each participant in Team-B gets B€*/3.
- If less than 3 participants accept the Proposer's new distribution, A€* and B€* are NOT implemented. Each participants get 0€.

2.3 Control Questions

We have designed some control questions to test your general understanding of the experiment. Please now submit your answers to the following questions on the computer.

1. Suppose that the envelope contains 90€. Which participants will NOT know this amount if no signal was sent?
(1=Proposer; 2=Members of Team-A; 3=Members of Team-B)

2. Suppose that the envelope contains 30€ and the Proposer plans to keep A€=15€ for Team-A. How much does the Proposer give to Team-B (B€)?
(1=€15; 2= €30; 3=€45)

3. Suppose that one Member of Team-A and two Members of Team-B accept the Proposer's distributions (initial or NEW). Will the distribution be implemented?
(1=Yes; 2=No)

4. Suppose that one Member of Team-A and one Member of Team-B accept the Proposer's distributions (initial or NEW). Will the distribution be implemented?
(1=Yes; 2=No)

5. How much does it cost to send a Signal?
(1=€0,00; 2=€0,50; 3=€1,00)

6. Suppose that the envelope contains 150€. Which participants will know this if a Signal is sent?
(1=Participants in Team-A only; 2=Participants in Team-B only; 3=all participants)

7. Suppose that the envelope contains 90€. Which participants will NOT know this if a Signal is NOT sent?
(1=Participants in Team-A only; 2=Participants in Team-B only; 3=all participants)`

8. How many Members of Team-A can send a Signal?
(1=both members; 2=only a randomly chosen member)

3 Design of the Experiment

During this experiment, you will interact with 5 other randomly selected participants in this room. At the start of the experiment the 6 participants will be randomly assigned into a Team (Team-A or Team-B) and a Role (Proposer, Member of Team-A or Member of Team-B). All participants will also be given an endowment of 1€.

The experiment consists of 4 stages. The decisions you will make in each stage depend on your Team and your Role. All of your decisions are anonymous, meaning that it is impossible for other participants to know what you did.

To help you better understand the experiment, we have organised the instructions as follows: Section 3.1 details the instructions for the Proposer, Section 3.2 for a Member of Team-A and Section 3.3 for a Member of Team-B.

After reading these instructions your Team and your Role will appear on your computer screen.

Please pay close attention to the instructions as the experiment will only consist of one round.

3.1 Instructions for the Proposer in Team-A

Stage 1

If you are a Proposer in Team-A your task in Stage 1 is to choose, *before knowing the actual amount of money in the envelope*, the amount of money you plan to keep for Team-A (A€) and the amount of money you plan to give to Team-B (B€).

You will be presented with the following two statements:

> "**I propose to keep_____(Your A€ Amount) Euro for Team-A**"
>
> **and**
>
> "**I propose to give_____(Your B€ Amount) Euro to Team-B**".

You will have to answer these two questions for each possible amount in the envelope: 30€, 90€, and 150€.

It is important that your amounts for A€ and B€ add up to the total amount in the envelope.

Stage 2

At the start of Stage 2, the computer randomly selects the actual amount of money in the envelope and all participants in Team-A are informed of this amount. Given this amount, we use your decisions (A€ and B€) from Stage 1. What happens next depends on whether a Signal is sent:

- If a Signal is sent, the Members of Team-B will also be informed of the amount in the envelope, so everyone will know A€ and B€. Everyone will then proceed to Stage 3.
- If no Signal is sent, the Members of Team-A will see A€ and B€, the Members of Team B will only see B€, and everyone will proceed to Stage 4.

Stage 3

You are in Stage 3 if a Signal has been sent, so the amount of money in the envelope has been revealed to all participants. You can revise your distribution plan. You choose the new amount of money you plan to keep for Team-A (A*€) and the new amount of money you plan to give to Team-B (B*€) for the actual amount of the money in the envelope that has just been revealed. You will be presented with the following two statements:

> "**I now propose to keep_____(Your A€* Amount) Euro for Team-A**"
>
> **and**
>
> "**I now propose to give_____(Your B€* Amount) Euro to Team-B**".

Stage 4

In Stage 4 the payoffs for all participants are determined. Your payoffs from this Task are computed as:

(1€ Endowment) + (Distributed Amount)

The Distributed Amount depends on the number of Members of Team-A and Team B who accepted your proposed distribution (A€ and B€, or A€* and B€*).

- If 3 or more accepted your distribution, your Distributed Amount is equal to A€/3 (if no Signal was sent) or A€*/3 (if a Signal was sent).
- If less than 3 accepted the distribution, your Distributed Amount is equal to 0€.

3.2 Instructions for the Members of Team-A

Stage 1

If you are a Member of Team-A you make two types of decisions in Stage 1: **(i)** your signalling decisions and **(ii)** your acceptance decisions.

Signalling decision

First, you choose, *before knowing the actual amount of money in the envelope*, the "conditions" under which you would like to send a Signal. You will be presented with the following statement:

"I would only like to send a Signal if the Proposer gives less than_____Euro to Team-B".

You will answer this question for each possible envelope amount: 30€, 90€, 150€. We will call your decision here your "Signal-Strategy".

You make your signalling decision before you know whether you are the randomly chosen member who can send a signal. **As such, you should always behave as if you are the only Member of Team-A who can send a signal.**

Please note the following:

- sending a Signal will cost you 0,50€
- put "0" as your Signal-Strategy if you NEVER want to send a Signal
- put "envelope amount" as your Signal-Strategy if you ALWAYS want to send a Signal

After both Members of Team-A have made their signalling decision, the computer will randomly choose one Member of Team-A and use her/his Signal-Strategy to determine whether a Signal is sent.

Table 1 gives an example that will help you understand the Signal-Strategy better. In this example, you have decided to always send a Signal if the envelope contains 30€ and to never send a Signal if the envelope contains 90€ or 150€.

Envelope Amount	30€	90€	150€
Your Signal-Strategy	30	0	0

Table 1. Example of Signal-Strategy

Table 2 is another example. Here, you have decided to send a Signal if

- the envelope contains 30€ and B€ < 10 (the amount given to Team-B is less than 10),
- the envelope contains 90€ and B€ < 30 (less than 30),
- the envelope contains 150€ and B€ < 34 (less than 34).

Envelope Amount	30€	90€	150€
Your Signal-Strategy	10	30	34

Table 2. Another Example of Signal-Strategy

Acceptance decision

After you have submitted your Signal-Strategy, you choose, *also before knowing the actual amount of money in the envelope*, the "conditions" under which you would reject the amount proposed by the Proposer to be kept for your Team (A€) if NO signal was sent. You will be presented with the following statement:

"I would only accept the Proposer's proposed distribution for Team-A (A€) if the Proposer keeps at least _____ Euro for Team A".

You will answer this question for each possible envelope amount: 30€, 90€, 150€. We will call your decision here your "Acceptance-Strategy for Team A".
Table 3 gives an example Acceptance-Strategy. In this example, you have decided to accept A€ if

- the envelope contains 30€ and A€ ≥ 10 (the amount kept for Team-A is at least 10),
- the envelope contains 90€ and A€ ≥ 30 (at least 30),
- the envelope contains 150€ and A€ ≥ 34 (at least 34).

Envelope Amount	30€	90€	150€
Your Acceptance Strategy	10	30	34

Table 3. Example of Acceptance-Strategy

Stage 2

At the start of Stage 2, the computer randomly selects the actual amount of money in the envelope and all participants in Team-A are informed of this amount. Given this amount, we use the decision from the Proposer to determine A€ and B€. Then, given the amount of money in the envelope and A€, we use Stage 1 decision from the randomly chosen Member of Team-A to determine Signal is sent. What happens next depends on whether the signal is sent.

- If the randomly chosen Member of Team-A sends a Signal, the Members of Team-B will also be informed of the amount in the envelope, so everyone will know A€ and B€. Everyone will then proceed to Stage 3.

- If no Signal is sent, the Members of Team-A will see A€ and B€, the Members of Team-B will only see B€ and everyone will proceed to Stage 4.

Stage 3

You are in Stage 3 if a Signal has been sent. The amount of money in the envelope is revealed to all participants. The Proposer has to make a new distribution plan and your task is to make again your acceptance decision by choosing the "condition" under which you would reject the new amount proposed by the Proposer to be kept for your Team (A*€) . You will be presented with the following statement:

"I would only accept the Proposer's new proposed distribution for Team A (A€*) if the Proposer keeps at least _____ Euro for Team A".

We will call your decision here your "New Acceptance-Strategy for Team A".

Here is an example. Suppose that the envelope contains €90 and your New Acceptance-Strategy is 60€. This means that you will only accept the Proposer's new distribution if A€*≥60 (if the Proposer keeps at least 60€ for Team-A).

Stage 4

In this stage the payoffs for all participants will be determined. Your payoffs from this Task are computed as:

(1€ Endowment) - (Cost for Signal, if any) + (Distributed Amount)

The Distributed Amount will depend on the number of Members of Team-A and Team-B who have accepted the distribution proposed by the Proposer (A€ and B€, or A€* and B€*).

- If 3 or more accepted the distribution, your Distributed Amount is equal to A€/3 (if no Signal was sent) or A€*/3 (if a Signal was sent).
- If less than 3 accepted the distribution, your Distributed Amount is equal to 0€.

It is important to note here that the Signalling cost is incurred as soon as your Signal-Strategy is greater than €B even when you are not the randomly chosen member. That is why it is important that you make your decisions as if you were the randomly selected Member of Team-A able to send the Signal.

3.3 Instructions for the Members of Team-B

Stage 1

If you are a Member of Team-B you choose in Stage 1 the "condition" over which you would accept the amount proposed by the Proposer for Team-B (B€) if NO Signal was sent. You will be presented with the following statement:

"I would only accept the Proposer's proposed distribution for Team-B (B€) if the Proposer gives at least ____ Euro to Team B".

We will call your decision here your "Acceptance-Strategy for Team B".

Here is an example that will help you understand it better. Suppose that no Signal was sent and that your Acceptance-Strategy is 30. This means that you will only accept the Proposer's distribution for Team-B if B€ ≥ 30 (if the Proposer gives at least 30€ to Team-B).

Stage 2

At the start of Stage 2, the computer randomly selects the actual amount of money in the envelope and all participants in Team-A are informed of this amount. Given this amount, we use the decision of the Proposer to determine A€ and B€. Then, given the amount of money in the envelope and A€, we use the decision of the randomly chosen Member of Team-A to determine whether a Signal is sent. What happens next depends on whether a Signal is sent:

- If a Signal is sent, you will be informed of the amount in the envelope and everyone will know A€ and B€. Everyone will then proceed to Stage 3.
- If no Signal is sent, you will only see B€ and everyone will proceed to Stage 4.

Stage 3

You are in Stage 3 if a Signal has been sent, so the amount of money in the envelope has been revealed to all participants. The Proposer decides on her/his new distributions and your task is to decide on the "condition" under which you would accept the new amount for Team-B (B€*). You will be presented with the following statement:

"I would only accept the Proposer's new proposed distribution for Team-B (B€*) if the Proposer gives at least _____ Euro to Team B".

We will call your decision here your "New Acceptance-Strategy for Team B". Here is an example that will help you understand it better. Suppose that your New Acceptance-Strategy is 60. This means that you will only accept the Proposer's new distribution for Team-B if B€*≥60 (if the Proposer gives at least 60€ to Team-B).

Stage 4

In this stage, the payoffs for all participants will be determined. Your payoffs from this Task are computed as:

(1€ Endowment) + (Distributed-Amount)

The Distributed Amount will depend on the number of Members of Team-A and Team-B who have accepted the distribution proposed by the Proposer (A€ and B€, or A€* and B€*).

- If 3 or more accepted the distribution, your Distributed Amount is equal to B€/3 (if no Signal was sent) or B€*/3 (if a Signal was sent).
- If less than 3 accepted the distribution, your Distributed Amount is equal to 0€.

Your Team and your Role now appears on your computer screen.

We have designed more control questions specific to your Role. Please read the control questions related to your Role below and enter your answers on the computer screen. You have a few minutes if you wish to re-read the instructions specific to your role in Section 3.

Control Questions for the Proposer of Team A

1. If you propose A€=30 for Team-A, and this proposal is implemented you receive_____.
   (1=€0; 2=€10; 3=€30)

2. If no Signal was sent in Stage 2, you_____ proceed to Stage 3 of the experiment.
   (1=Will; 2=Will NOT)

3. If a Signal was sent in Stage 2, you _____ proceed to Stage 3 of the experiment.
   (1=Will; 2=Will NOT)

4. If Team-B was sent a Signal, you can make a second proposal on the distributions for Team-A and Team-B. Suppose that you now propose to keep A€*=60 for Team A and the proposal is implemented. You will receive _____.
   (1=€0; 2=€60; 3=€20)

5. If Team B was sent a Signal, you can make a second proposal on the distributions for Team-A and Team-B. Suppose that you now propose to keep A€*=60 for Team A and the proposal is NOT implemented. You will receive _____.
   (1=€0; 2=€60; 3=€20)

Control Questions for the Members of Team A

1. A Signal-Strategy of 0 in Stage 1 implies that_____.
   (1=You will always send a Signal; 2=You send a Signal only if Team B receives less than 20€; 3=You will never send a Signal)

2. A Signal-Strategy of 30 in Stage 1 implies that_____.
   (1=You will send a Signal if more than 30 is proposed for the entire Team-B; 2=You will send a Signal if less than 30 is proposed for the entire Team-B; 3=You will send a Signal if less than 30 is proposed for each participant in Team B)

3. An Acceptance-Strategy of 15 in Stage 1 implies that_____.
   (1=You will accept any proposal that keeps at least 15 for the entire Team-A; 2=You will accept any proposal that keeps at most 15 for the entire Team-A; 3=You will accept any proposal that keeps at least 15 for each participant in Team-A)

4. If no Signal was sent by the randomly chosen Member of Team-A, you _____ proceed to Stage 3 of the experiment.
   (1=Will, 2=Will NOT)

5. If a Signal was sent by the randomly chosen Member of Team-A, all participants in Team A and all participants in Team-B _____ know in Stage 2 the amount of money in the envelope and the proposed amounts by the Proposer.
   (1=Will, 2=Will NOT)

Control Questions for the Members of Team B

1. If your Acceptance-Strategy is 25, it means that _____ .
   (1=You will accept any proposal that gives at least 25 to the entire Team-B; 2=You will accept any proposal that gives at most 25 to the entire Team-B; 3=You will accept any proposal that gives at least 25 to each participant in Team-B)

2. When you choose your Acceptance-Strategy in Stage 1___.
   (1=You know the amount of money in the envelope; 2=You do not know the amount of money in the envelope; 3=In some cases you may know the amount of money in the envelope)

3. When you decide about your New Acceptance-Strategy in Stage 3, you _____ the amount of money in the envelope and the previously proposed amounts for Team A and Team B.
   (1=Will know; 2=Will Not know)

Important information

All of what we have just described will happen only once and there will be no trial or practice, so it is important that you understand it well. Please raise your hand if there are any questions and the experimenter will answer your questions privately.

After everyone in the room has completed Stage 1, we will pause the First Task and start the Second Task. Once the Second Task is finished, we will continue and finish the First Task. Then, we will determine your payment as we have already explained.

Second Task

We will now explain the second task of the experiment. For this task, we will present you with 'situations'. Each situation features a particular action that either a Proposer or a Member of Team A could take in the First task. You will be asked to evaluate whether you consider taking this action to be 'socially appropriate' and 'consistent with moral or proper social behaviour' or 'socially inappropriate' and 'inconsistent with moral or proper social behaviour'.

By socially appropriate, we mean actions that most people agree to be the 'correct' or 'ethical' thing to do. Another way to think about what we mean is that if an 'Individual' were to take a socially inappropriate action, then someone else might be angry at the Individual for doing so. In each of your responses, we would like you to answer as truthfully as possible, based on your opinions of what most people think is socially appropriate or socially inappropriate behaviour.

To give you an idea of how the experiment will proceed, we will go through an example and show you how you will indicate your responses. This example is not related to today's experiment and is merely here for illustration purposes.

Example situation

An Individual is at a local coffee shop near university. While being there, the Individual notices that someone has left a wallet on one of the tables. The Individual must decide what to do. For example, the Individual can take the wallet, ask others nearby if the wallet belongs to them, leave the wallet where it is, or give the wallet to the shop manager.

Let us focus on the action 'leave the wallet'. This action can be taken in different contexts. The table below presents three possible contexts: there are many people around, there is no one around, or the Individual has already asked people around and the wallet is not theirs. You will see tables similar to the below example on your computer screen during the experiment.

Individual's action	when...	Very socially inappropriate	Somewhat socially inappropriate	Somewhat socially appropriate	Very socially appropriate
Leave the wallet	there is no one around	○	○	○	○
Leave the wallet	there are many people around	○	○	○	○
Leave the wallet	the wallet belongs to no one around	○	○	○	○

If this were one of the situations for this study, you would consider the action 'leave the wallet' in each of the possible contexts above and, for that context, indicate the extent to which you believe taking that action would be 'socially appropriate' and 'consistent with moral or proper social behaviour' or 'socially inappropriate' and 'inconsistent with moral or proper social behaviour' by double-clicking with your mouse on the corresponding button. **Recall that by socially appropriate we mean behaviour that most people agree is the 'correct' or 'ethical' thing to do.**

For example, suppose you thought that most people find leaving the wallet somewhat socially appropriate when there are people around, somewhat socially inappropriate when there is no one around and very socially inappropriate when others around have been asked and the wallet is not theirs. Then you would indicate your responses as follows:

Individual's action	when...	Very socially inappropriate	Somewhat socially inappropriate	Somewhat socially appropriate	Very socially appropriate
Leave the wallet	there is no one around	○	●	○	○
Leave the wallet	there are many people around	○	○	●	○
Leave the wallet	the wallet belongs to no one around	●	○	○	○

Next, we could focus on the action 'take the wallet' and ask you to evaluate this action with the same three contexts. This would form a new situation.

Are there any questions about this example?

Your decisions in the Second task

You will face 21 situations in total. A given situation will feature an action available to a Proposer or a Member of Team-A. The Role of the participant will be clearly indicated at the top of the screen. For situations featuring a Proposer, you will be asked to evaluate the social appropriateness of some offers that the Proposer can make to Team-B. For those featuring a Member of Team-A, you will evaluate the social appropriateness of not sending a Signal. Note that you will not see actions available to a Member of Team-B.

You will see the situations one by one on your computer screen. The different envelope amounts will be the contexts in which you will evaluate the actions of the Proposer or of a Member of Team-A. You will read a description of the situation and indicate whether the action is socially appropriate or socially inappropriate depending on the context. You will indicate your responses using a table similar to the one shown above in the wallet example situation. The table will also be presented on your computer screen and you will have to double-click on the radio buttons to enter your answer. Note that, once you move to the next situation, you cannot go back to the previous one to change your responses.

Your cash earnings if the second task is selected for payment

As we mentioned in the earlier instructions, you will only be paid for one task: the First Task or the Second Task. At the end of the experiment the computer will randomly determine which task you will be paid for. If the computer determines that you are to be paid for the Second Task, we will calculate your payment as follows.

The computer will first select one of the 21 situations. From the selection situation, the computer will then randomly select one of the possible contexts in which the Individual can take the action featured in the situation. Thus, the computer will select both a situation and a context at random.

For the selected situation and context, we will determine which response was chosen most frequently by the people in this room today.

- If your response is the same as the most frequent response chosen by all people, then you will receive 20€. This amount will be paid to you at the end of the experiment, in cash, and in addition to the fixed payment of 4€.
- If your response is different than the most frequent response chosen by all people, then you will only receive the fixed payment of 4€.

To better understand the payment from the second task, consider the wallet example above. Suppose that the computer randomly chooses the context 'there is no one around' and the most frequent response was 'somewhat socially inappropriate'. In this case,

- you would receive 20€ , in addition to the fixed payment of 4€, if your response was 'somewhat socially inappropriate' .
- you would receive only the fixed payment of 4€ if your response was different from 'somewhat socially inappropriate'.

*   *   *

The Second Task will now begin shortly. Please raise your hand if you have any questions and an experimenter will come to your desk to answer it.

References

Benjamini, Yoav and Yosef Hochberg (1995), "Controlling the false discovery rate: A practical and powerful approach to multiple testing." *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 289–300.

Cohen, Jacob (1988), *Statistical Power Analysis for the Behavioural Sciences*, second edition. Lawrence Erlbaum Associates.

Fan, Chunpeng and Donghui Zhang (2012), "A note on power and sample size calculations for the Kruskal-Wallis test for ordered categorical data." *Journal of Biopharmaceutical Statistics*, 22, 1162–1173.

Fan, Chunpeng, Donghui Zhang, and Cun Hui Zhang (2011), "On sample size of the Kruskal-Wallis test with application to a mouse peritoneal cavity study." *Biometrics*, 67, 213–224.

Faul, Franz, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner (2007), "G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences." *Behavior Research Methods*, 39, 175–191.

Kupper, Lawrence L. and Kerry B. Hafner (1989), "How appropriate are population sample size formulas?" *The American Statistician*, 43, 101–105.

Lachenbruch, Peter A. and Philip J. Clements (1991), "ANOVA, Kruskal-Wallis, normal scores and unequal variance." *Communications in Statistics - Theory and Methods*, 20, 107–126.

Lehr, Robert (1992), "Sixteen S-squared over D-squared: A relation for crude sample size estimates." *Statistics in Medicine*, 11, 1099–1102.

Murphy, Ryan O., Kurt A. Ackermann, and Michel J. J. Handgraaf (2011), "Measuring social value orientation." *Judgment and Decision Making*, 6, 771–781.

Sawilowsky, Shlomo S. (2009), "New effect size rules of thumb." *Journal of Modern Applied Statistical Methods*, 8, 467–474.