Are groups really more dishonest than individuals?

Geoffrey Castillo* Lawrence Choo† Veronika Grimm†‡

28th April 2020

Abstract

A common finding of the literature on dishonesty is that groups are more dishonest than individuals. We revisit this finding by replacing the experimenter, implicitly hurt by subjects' dishonesty, by an explicit third-party: a local charity. With this explicit third-party, we do not find groups to be more dishonest than individuals. Instead, we find that groups can even help moderate the extent of dishonesty.

JEL-classification: *C91, C92, D71.*
Keywords: Dishonesty, Group decisions, Communication, Social norms.

*Vienna Centre for Experimental Economics, University of Vienna, Austria.

†Department of Economics, University of Erlangen-Nuremberg. Corresponding author: lawrence.cy.choo@fau.de

‡We are grateful to the student assistants of the LERN laboratory who helped us to conduct the sessions. We also thank numerous seminar and conference participants for helpful comments.

1 Introduction

Recently, much has been written about dishonesty in the laboratory (Fischbacher and Föllmi-Heusi, 2013; Gneezy, 2005; Mazar et al., 2008). One of the common findings in this literature is that groups are more dishonest than individuals (Kocher et al., 2018).[1]

In most of these studies the extra earnings subjects receive by acting dishonestly come from the experimenter's budget, so dishonesty is understood as an action that implicitly only harms the experimenter. Outside the laboratory, however, dishonesty often harms explicit third-parties, such as when financial advisers embezzle the funds of investors. Further, if groups were really more dishonest than individuals, decision-making in groups would not be promoted so widely.

In our experiment, to make the third-party harmed by subjects' dishonesty more salient, we replace the experimenter by a local charity. We observe that, as soon as we introduce this charity, groups stop being more dishonest than individuals. In fact we find that groups even help reduce the extent of the dishonesty. One of the reasons is that, with the charity, participants stop sharing opinions in favour of dishonesty.

We base our experimental design on Kocher et al. (2018), itself a variant of the Fischbacher and Föllmi-Heusi (2013) die-rolling task.[2] In this task subjects roll a die and must truthfully report the number they see, but their payoff only depends on their report, not on the number seen. The higher the number reported, the more money is implicitly taken from the experimenter. One of the innovations of Kocher et al. (2018) is to use a video of a die roll; since the experimenter knows which videos subjects see it is possible to study dishonesty at the individual level. They compare individual and group behaviour and allow group members to chat to coordinate on a number to report.

In our *Base* treatments we replicate their experiment and find the same result: while individuals and groups both report dishonestly, groups are much more dishonest than individuals.

We then introduce two extensions: the *Charity* and *CharityR* treatments. The monetary incentives remain identical across all treatments. In the *Charity* and *CharityR* treatments, however, subjects' decisions also affect the amount of money that the experimenter will anonymously donate to a local charity. In doing so, we

[1] See also Barr and Michailidou (2017); Conrads et al. (2013); Gino et al. (2013); Chytilová and Korbel (2014); Muehlheusser et al. (2015); Weisel and Shalvi (2015); Soraperra et al. (2017).

[2] See Abeler et al. (2019) for an extensive review of the literature using this task.

create an environment where dishonest behaviour explicitly harms a third-party. The *CharityR* treatments differ from the *Charity* only in that subjects' decisions are revealed to the charity.

Our first main result is that, with the charity, groups are no more likely to act dishonestly than individuals. In fact, if we look at the intensive margin, we find that dishonest groups cheat to a lesser extent than dishonest individuals. Our second main result is that group dishonesty can be contagious: subjects who were previously members of a dishonest group are more likely to act dishonestly in the future.

To understand why the introduction of the charity has such a clear impact, we study how individuals and groups adjust their decisions when we introduce the charity. We find that individual dishonesty is stable across treatments and does not respond to the introduction of the charity. This finding is consistent with Fischbacher and Föllmi-Heusi (2013) who observe that changing the consequences of dishonesty has little influence on individual dishonesty.

On the other hand, we find that groups are more dishonest in the *Base* treatments than in the *Charity* and *CharityR* treatments. To explain the difference between individuals and groups we look at the chat data and find that subjects chatting in groups in the *Charity* and *CharityR* treatments make lesser arguments for dishonesty than those in the *Base* treatment.

Since monetary incentives are constant across treatments, they alone cannot explain these finding. Different other-regarding preferences toward the experimenter and toward the charity could explain different baseline level of dishonesty but would not be enough to explain the difference between the individual and group treatments. Instead the evidence points toward social norms and how they are shared in the chat. In the *Base* treatments, it is easy for dishonest group members to express preferences for dishonesty. Therefore, those who would have acted honestly if they had made the decision alone are now exposed to arguments in favour of dishonesty and may acquiesce to them if it helps the group reach a decision quicker. By contrast, in the *Charity* and *CharityR* treatments it is more difficult to express preferences for dishonesty since doing so signals a disdain for the charity and so taints one's social image. The move toward more honesty can even be reinforced by moral reminders (e.g., Pruckner and Sausgruber, 2013) from honest group members.

We also contribute to the literature on dishonesty by showing that, while we should strive for abstractness and neutrality when running experiments, it is sometimes appropriate to bring more elements from real-world situations into

Table 1: Payoffs to subjects and donations to the charity.

Die	⚀	⚁	⚂	⚃	⚄	⚅
Points	0	1	2	3	4	5
Payoff to subject	0 €	2 €	4 €	6 €	8 €	10 €
Donation to charity (*Charity* and *CharityR* treatments only)	10 €	8 €	6 €	4 €	2 €	0 €

the laboratory. Based on the existing evidence one could conclude that decision-making in groups is detrimental because it leads to higher dishonesty. In few situations, however, is the person or the organisation harmed by the dishonesty completely absent. When employees steal from their company, or when citizens cheat on their tax declaration, there is a clear victim. We show that, as soon as we reintroduce a victim, the conclusions change and groups become no more dishonest than individuals—if anything they even regulate the extent of the dishonesty.

In the next Section we present our experiment design. Section 3 reports our results. Finally, Section 4 concludes.

2 Experimental design

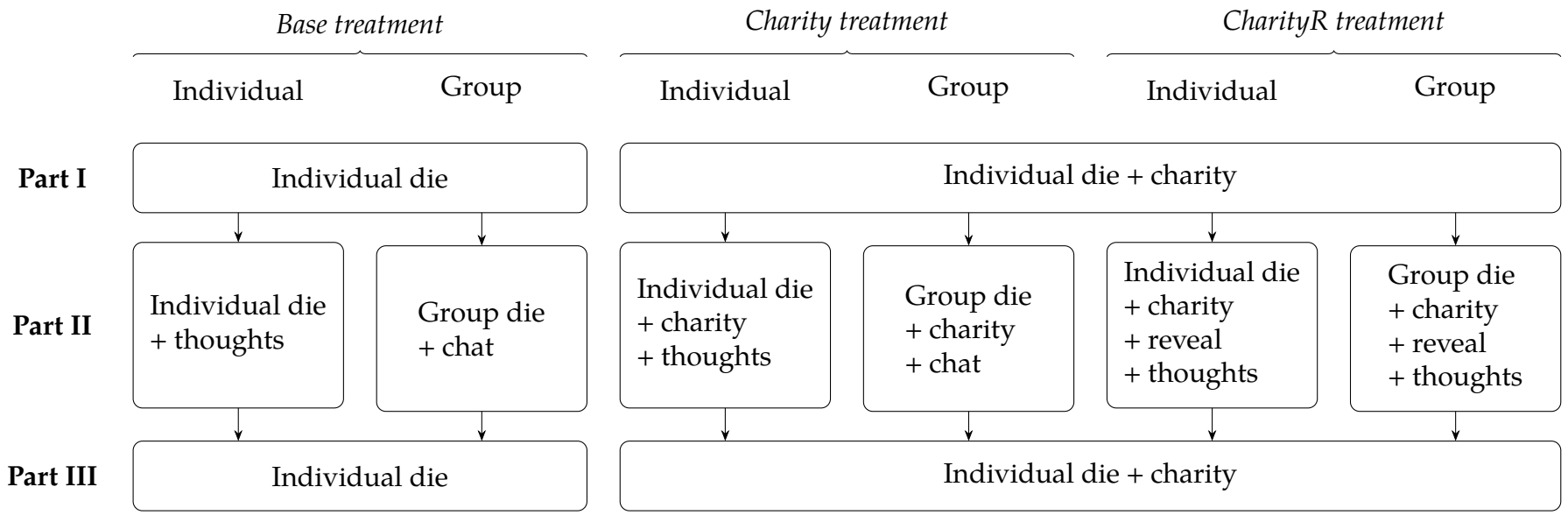
2.1 Treatments

Our experiment involves six treatments, grouped into three categories: *Base*, *Charity* and *CharityR*. Figure 1, which we will use throughout, summarises our design.

***Base* treatments.** To the left of the Figure are our *Base* individual (*I-Base*) and group (*G-Base*) treatments, which replicate the "Individual" and "GroupPC" treatments of Kocher et al. (2018).

There are three Parts to each treatment. In Parts I and III, subjects watch a video of a die roll and are asked to *truthfully* report the result of the roll which can be ⚁, ⚂, ⚃, ⚄, ⚅ or ⚀ with equal probability. The top part of Table 1 shows that the reported die roll number is converted into points so that ⚀ corresponds to 0 points, ⚁ to 1 point, ⚂ to 2 points, and so on; and that each point is worth 2 €.

What differs between the *I-Base* and *G-Base* treatments is the second part. In



Notes. **+ charity**: subjects' decisions also affect the donations the experimenter makes to a local charity.
+ thought: subjects can write their thoughts about their decision (for maximum 5 minutes).
+ chat: subjects of the same group can chat (for maximum 5 minutes).
+ reveal: subjects' decisions are revealed to the charity.

Figure 1: Summary of the experimental design.

the *I-Base* treatment, subjects do the same **i**ndividual task as in Parts I and III. In the *G-Base* treatment, however, subjects form **g**roups of three, all observe the same die roll video, and then independently report what they observed. Group members each receive the points shown in Table 1 if they all reported the same number, and 0 points otherwise. For example, each group member receives 3 points (6 €) if all three report ⚂. On the other hand, each group member receives 0 points (0 €) if they fail to report the same number.

To coordinate, group members can chat anonymously for up to 5 minutes through the experimental software. To keep the treatments comparable, subjects in Part II of the *I-Base* treatment are given 5 minutes to, if they wish, type their thoughts about the experiment.

***Charity* treatments.** The *Charity* individual (*I-Charity*) and group *(G-Charity)* treatments are identical to their respective *Base* treatments, except that they make explicit that a passive third-party, a local charity, is hurt by the report of a large number: each extra Euro given to the subject as a result of their report is not given to a local charity. For example as shown on Table 1, if a subject reports ⚂, the subject gets 6 € and the charity gets 4 €. In Part II of the *G-Charity* treatment, group members and the charity receive 0 € if the group members do not report the same die number.

***CharityR* treatments.** In the *CharityR* treatments, we further revealed to the charity the die number observed and reported in Part II of the individual (*I-CharityR*) and group (*G-CharityR*) treatments, as well as the written thoughts (*I-CharityR*) and chat logs (*G-CharityR*). All of this information was anonymous. Parts I and III were unchanged.

2.2 Procedures

When the experiment started, subjects were told that the experiment was made of three independent parts. At the beginning of each part they received specific instructions and answered some control questions where we made sure that they understood how their report could affect them and, if applicable, the charity. In particular the control questions made clear that, even if their stated task was to truthfully report the die roll, their payoffs and the donations to the charity only depended on their report.

We used a small local charity whose members perform as clowns in nearby hospitals to entertain sick children. In the initial instructions of the *Charity* and

CharityR treatments, we gave subjects a broad description of the charity; they were then given 5 minutes to visit its website and Facebook page to learn about its activities. In the *CharityR* treatments, in which we revealed to the charity the subjects' decisions in Part II, we also gave subjects a sample of the information we would send to the charity via email. Subjects also knew they would be included in blind carbon copy to the email.

The experiment was conducted at the LERN of the University of Erlangen-Nuremberg between June 2018 and March 2019. Subjects were recruited via ORSEE (Greiner, 2015). We recruited 30 subjects for each of the individual treatments and 90 subjects for each of the group treatments, ensuring the same number of independent observations in each case. The experiment was programmed with zTree (Fischbacher, 2007).[3] The instructions are presented in Appendix D.2.

Following Kocher et al. (2018) we randomly generated in the first session one sequence of die rolls for each group and used the same sequences in the next sessions. We can therefore directly compare the die reports across treatments. This procedure also ensured that subjects of the same group of the *G-Base*, *G-Charity* and *G-CharityR* treatments also observed the same die roll in Parts I and III.

At the end of each session, one of the three parts was selected at random to determine subjects' earnings and the donations to the charity. To convince subjects that the charity and the donations were real, we made the anonymous donations immediately after the end of the tasks from within the laboratory using a laptop, and subjects were actively encouraged to monitor the process. Subjects knew this when they started the experiment.

Each experimental session lasted approximately 60 minutes. The mean earnings for subjects in the *Base*, *Charity* and *CharityR* treatments, including a show-up payment of 4 €, are 13.45 €, 12.06 € and 10.84 €, respectively.

2.3 What can we expect?

Since our *Base* treatments replicate Kocher et al. (2018), and since they observed more dishonesty with groups than with individuals, we can expect that too. In the *Charity* and *CharityR* treatments, however, dishonesty affects a named third-party, the charity. We purposely use a charity and not another experimental subject to bring pre-existing "homegrown norms" (Schram and Charness, 2015) into the laboratory. Changes in psychological disutility stemming from social

[3]We are grateful to Martin Kocher, Simeon Schudy and Lisa Spantig for making their zTree code available to us, which we adapted to our experiment.

image concerns, and in social norms potentially shared in group communication, can shift the results across treatments.

In the individual treatments, subjects who observe less than 5 points and who contemplate reporting more points than observed face a trade-off between monetary gain and psychological disutility. The monetary gain is constant across treatments and only the psychological disutility varies in Part II of the respective treatments; it is arguably low in the *I-Base* treatment, moderate in the *I-Charity* treatment, and, since the charity is informed of the subjects' decisions, high in the *I-CharityR* treatment.

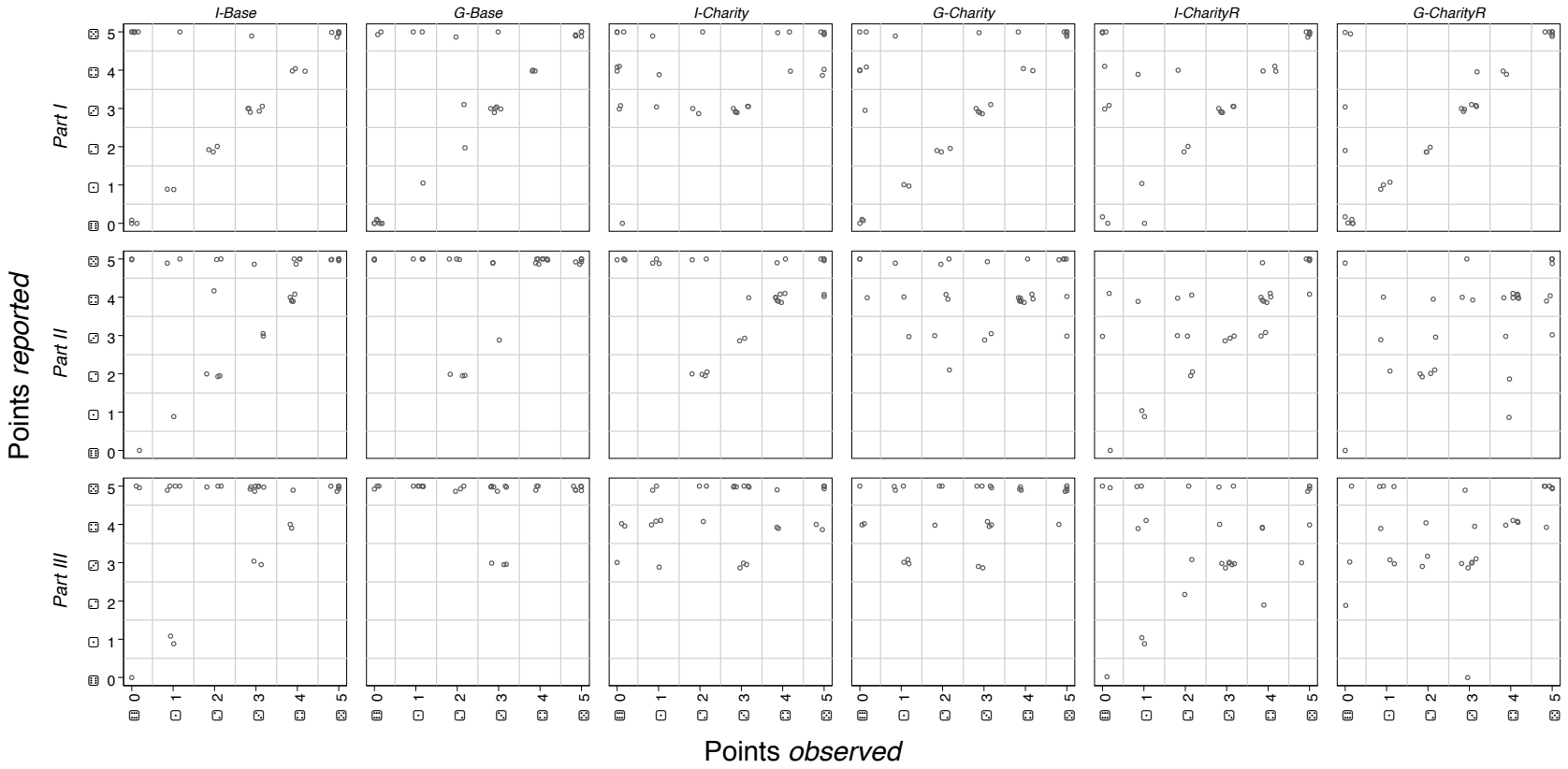
The group treatments (*G-Base*, *G-Charity*, *G-CharityR*) differ from the individual treatments (*I-Base*, *I-Charity*, *I-CharityR*) in the necessity for group members to coordinate their decisions in Part II. When studied as a coordination game, there are multiple equilibria, some of which can be more focal than others. For groups in the *G-Base* treatment who observe less than 5 points, the equilibria where all report honestly and all report the maximum possible 5 points are intuitively focal. This is less clear in the *G-Charity* and *G-CharityR* treatments where groups may still be dishonest but report less than 5 points to ensure that the charity receives something—the equilibria in between the extrema where all report honestly and all report the maximum points can also be focal. Even if all group members privately prefer to report the maximum possible points, social norms can make it difficult for any member to express such preferences. In this case, the group's decision may be honest even when the majority of group members would have over-reported when making the decision alone. Hence, groups may be more honest than individuals in the *Charity* and *CharityR* treatments.

3 Results

Figure 2 shows, for all treatments and all parts, the numbers reported (y-axis) as a function of the numbers seen (x-axis). We organise the Figure in terms of points but also indicate the corresponding die numbers; the axis thus start with ⚃ = 0 points. In the group treatments, for ease of comparability we follow Kocher et al. (2018) and look at the median of the points reported in each group.

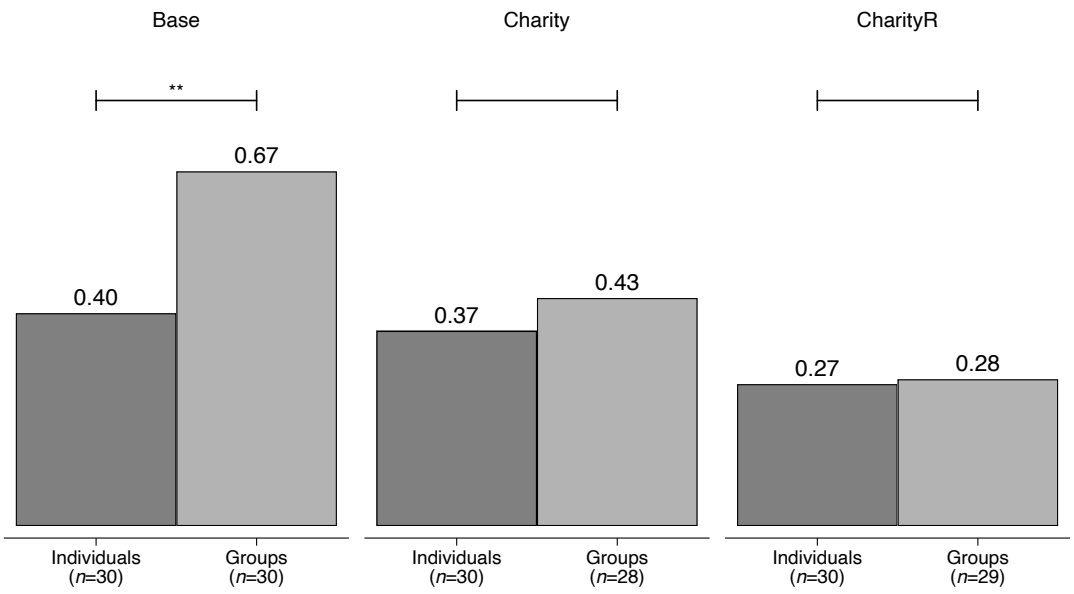
3.1 Groups are not always more dishonest than individuals

For the time being, we focus on Part II. Here, 100%, 93% and 97% of the groups in the *Base*, *Charity* and *CharityR* treatments reported the same number—we omit



Note. In the individual treatments, each dot represents the point observed and reported by a participant. In the group treatments, each dot represent the point observed (all group members observe the same points in Parts I, II and III) and the median point reported by a group.

Figure 2: Points observed and points reported across treatments and parts (jittered).



Note. Pearson's $\chi^2(1)$ tests; ***, ** and * denote $p < 0.01$, $p < 0.05$ and $p < 0.10$.

Figure 3: Proportion of individuals and groups that over-report in Part II.

observations where members fail to coordinate on their report in Part II. We say that individuals or groups *over-report* if their report gets them more points than if they had reported honestly.[4] Figure 3 shows the proportion of individuals and groups who over-report in Part II of each treatment.

As in Kocher et al. (2018), we find that in the *Base* treatments groups over-report more than individuals: 67% of groups over-report compared to 40% of individuals ($\chi^2(1) = 4.286$, $p = 0.038$; all statistical tests are two-sided). Note that our die sequences are different from the ones used by Kocher et al. (2018) (see Appendix D.1 for a comparison), which further strengthens the replicability of their findings:

Observation (Replication of Kocher et al., 2018)**.** *Groups over-report more frequently than individuals when over-reporting harms the experimenter.*

In the treatments involving the charity, however, we do not find that groups over-report more frequently than individuals. Figure 3 shows that the proportions of over-reporting groups and individuals in the *Charity* treatments are 43%

[4]All subjects made individual decisions in Part I. Here, we do not find any significant differences in the proportion of individual and group treatment subjects who over-report in the *Base* ($\chi^2(1) = 0.054$, $p = 0.815$) and *CharityR* ($\chi^2(1) = 0.742$, $p = 0.389$) treatments. There is, however, some differences in the *Charity* ($\chi^2(1) = 2.963$, $p = 0.085$) treatments, where individuals are slightly more dishonest than groups. For this reason we will also report regressions controlling for behaviour in Part I.

and 37% ($\chi^2(1) = 0.231$, $p = 0.630$), respectively and in the *CharityR* treatment are 28% and 27% ($\chi^2(1) = 0.006$, $p = 0.937$), respectively. The differences remain insignificant ($\chi^2(1) \leq 0.348$, $p \geq 0.555$) even if we exclude instances where subjects observed a die number corresponding to 5 points, for which they could not have over-reported.

If we look, not at the proportion of individuals or groups who over-report, but at the reported points, a similar picture emerges: groups report significantly higher points than individuals in the *Base* treatments (Mann-Whitney $U = 325$, $p = 0.019$), but there are no significant differences in the reported points of individuals and groups in the *Charity* ($U = 407$, $p = 0.827$) and *CharityR* ($U = 419$, $p = 0.806$) treatments. These observations lead us to our first result:

Result 1. *There is no significant difference in over-reporting between groups and individuals when over-reporting harms the charity. This holds even when the charity is informed.*

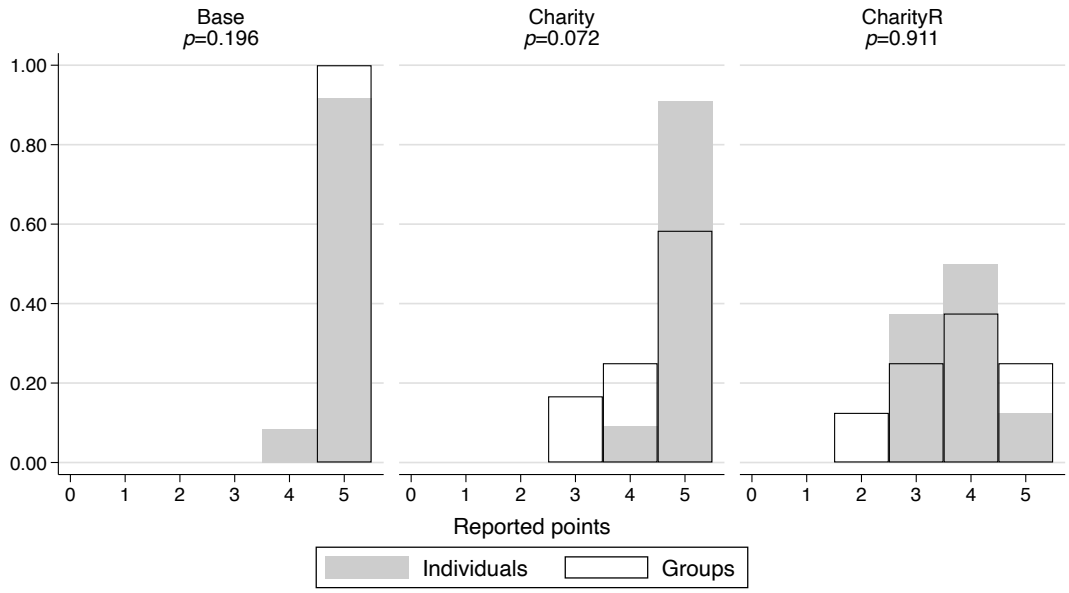
To give more meaning to these numbers, we can look at the extra revenue individuals and groups extract by over-reporting. In Part II of the *Base* treatments, individuals on average claim 2.14 € more than if they had reported honestly, and group members, 3.26 €. In line with the previous findings, the difference shrinks in the *Charity* treatments: individuals claim 2.26 € more, but group members now claim only 1.80 € more. In the *CharityR* treatments, the numbers for individuals and group members are even lower, respectively 0.94 € and 0.26 €. More details can be found in Appendix A.

It is clear from Figure 3 that groups not over-reporting more than individuals with the charity is not driven by individuals increasing their over-reporting: the proportion of over-reporting individuals in the *Base* treatment is not significantly different from those in the *Charity* ($\chi^2(1) = 0.070$, $p = 0.791$) and *CharityR* ($\chi^2(1) = 1.200$, $p = 0.273$) treatments.[5]

Instead, Figure 3 shows that the proportion of over-reporting groups in Part II falls sharply when we introduce the charity: the proportion of over-reporting groups in the *Base* treatment is significantly higher than those in the *Charity* ($\chi^2(1) = 3.320$, $p = 0.068$) and *CharityR* ($\chi^2(1) = 9.032$, $p = 0.003$) treatments.

Also, revealing subjects' behaviour to the charity has little influence: the proportion of over-reporting individuals in Part II of the *I-Charity* and *I-CharityR* treatments are not significantly different ($\chi^2(1) = 0.693$, $p = 0.405$). Similarly, the

[5]We observe the same in Part I: the proportion of over-reporting individuals in the *Base* treatments is 28%, compared to 32% for the *Charity* and *CharityR* treatments ($\chi^2(1) = 0.527$, $p = 0.468$), which we pool since at this point subjects faced the exact same task.



Note. p-values from Mann-Whitney U tests comparing the reported points of over-reporting individuals and groups.

Figure 4: Histogram of reported points by over-reporting individuals and groups in Part II.

proportion of over-reporting groups in Part II of the *G-Charity* and *G-CharityR* treatments are not significantly different ($\chi^2(1) = 1.458$, $p = 0.227$).

These observations lead us to the next result:

Result 2. *The proportion of over-reporting groups, but not of over-reporting individuals, falls when the charity is introduced. Revealing information to the charity, however, has no effect.*

Focus now on these over-reporting individuals and groups. There are no significant differences in the points they observed in the *Base* (Mann-Whitney $U = 107$, $p = 0.583$), *Charity* ($U = 65$, $p = 0.924$) and *CharityR* ($U = 30$, $p = 0.787$) treatments, so we can concentrate on the points they report.

Figure 4 shows the histogram of the reported points of these over-reporting individuals or groups. In the *Base* treatments, they report the same points ($U = 110$, $p = 0.196$): almost always the maximum possible. In the *Charity* treatments, however, over-reporting groups report significantly lower points ($U = 44$, $p = 0.072$) than over-reporting individuals. The difference remains significant ($U = 32$, $p = 0.089$) even after excluding instances where over-reporting individuals or groups observe 4 points, for which they could only report 5 points. In the

CharityR treatments the effect is in the same direction but it is not significant ($U = 31$, $p = 0.911$). Hence our third result:

Result 3. *Over-reporting groups over-report less than over-reporting individuals when group members know that over-reporting harms the charity.*

In Appendix B.1, we report regressions corresponding to Results 1, 2 and 3, where we further control for the die roll observed in Part II, the behaviour in Part I, and gender. All the results carry through.

In Figure 4 we can also see that, when moving from the *Base* treatment to the *Charity* treatment, the points reported by over-reporting groups fall sharply (Mann-Whitney $U = 70$, $p < 0.001$) while those reported by over-reporting individuals remain stable ($U = 66$, $p = 0.949$). When moving from the *Charity* treatment to the *CharityR* treatment, however, it is now the points reported by over-reporting individuals that fall ($U = 8$, $p < 0.001$) while those reported by over-reporting groups remain the same ($U = 30$, $p = 0.127$). One possibility is that individuals in the *CharityR* treatment start experiencing what groups already experienced in the *Charity* treatment: having your opinions and decisions revealed, be it to other group members or to the charity itself.

Finally, note that we observe no under-reporting—reporting fewer points than observed—in the *Base* treatments. We observe some under-reporting in the treatments involving the charity, where under-reporting would allow subjects to allocate to the charity more than if they had reported honestly, but there are no differences between individuals and groups: the proportion of individuals and groups who under-report are 6.67% and 7.14% ($\chi^2(1) = 0.005$, $p = 0.943$), respectively, in the *Charity* treatments and 10.00% and 20.69% ($\chi^2(1) = 1.303$, $p = 0.254$), respectively, in the *CharityR* treatments.

3.2 Why are groups not more dishonest than individuals when the charity is harmed? Analysis of the chat data

We have established that groups stop to over-report more than individuals when we introduce a charity. To better understand why, we look at the chat data. For each group, we compute the proportion of statements for dishonesty used in the chat.[6] As can be expected, there is a positive and significant correlation between

[6]The content analysis of the chat data was performed independently by a third-party. We looked for statements in which subjects made recommendations for the group decision or for statements in which subjects agreed with a previous recommendation. We then looked at the proportion of statements for which the recommended number of points was below the one observed. We provide a more detailed explanation in Appendix C.

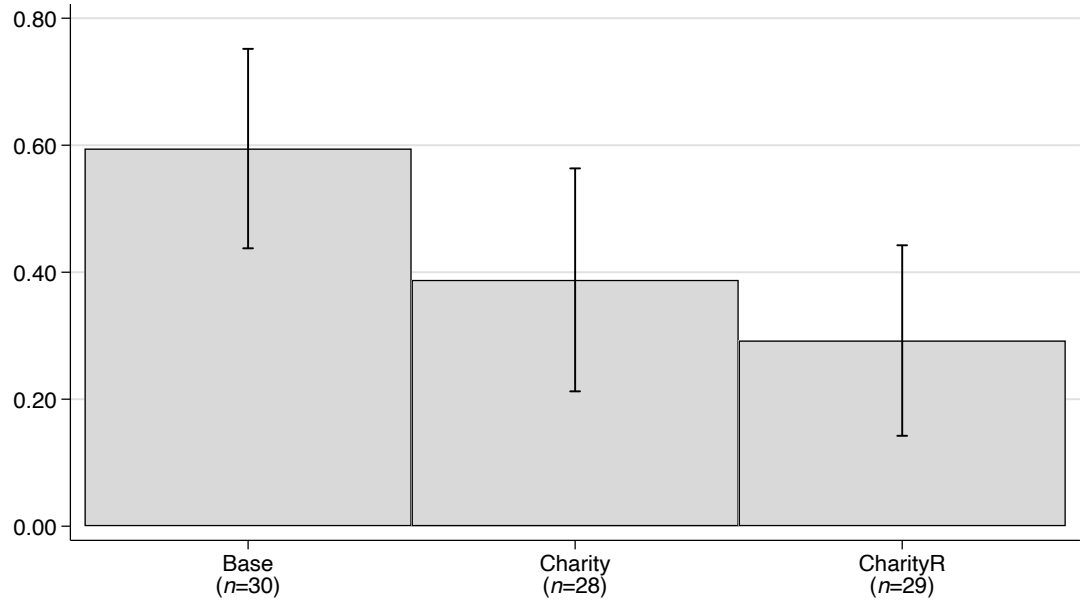


Figure 5: Mean and 95% confidence interval of the proportion of statements for dishonesty in Part II of the group treatments.

the proportion of statements for dishonesty and over-reporting for groups in the *Base* (Spearman's $\rho = 0.841$, $p < 0.001$), *Charity* ($\rho = 0.915$, $p < 0.001$) and *CharityR* ($\rho = 0.850$, $p < 0.001$) treatments.

Figure 5 details the mean and 95% confidence interval of the proportion of statements for dishonesty for groups in the different treatments. We observe that this proportion is higher in the *Base* treatment relative to the *Charity* (Mann-Whitney $U = 329$, $p = 0.142$) and *CharityR* ($U = 274$, $p = 0.011$) treatments. In Table 2, we use a fractional logit model (e.g., Papke and Wooldridge, 1996, 2008) to further control for the die number observed in Part II as well as the proportion of group members who over-reported in Part I. The regressions show that the proportion of statements for dishonesty is significantly higher for groups in the *Base* treatment than in the *Charity* ($p \leq 0.035$) and in *CharityR* ($p < 0.001$) treatments. There are no significant differences between groups in the *Charity* and the *CharityR* treatments ($p \geq 0.331$). This leads us to the next result:

Result 4. *Group members make significantly less arguments for dishonesty when they know that dishonesty harms the charity.*

These findings suggest that individuals' preference for over-reporting is relatively insensitive to the psychological disutility associated with dishonest behaviour. Groups, on the other hand, over-report more than individuals in the *Base* treatments but not the *Charity* and *CharityR* treatments because group members

Table 2: Fractional Logit regression estimates: How the proportion of dishonesty for over-reporting differs across groups in the group treatments.

Dependent Variable: Proportion of statements for dishonesty		
Regression	(1)	(2)
Reference group: Groups in the *G-Base* treatment.		
G-Charity	-1.09**	-1.14**
	(0.51)	(0.56)
G-CharityR	-1.60***	-1.54***
	(0.50)	(0.57)
Points observed (Part II)		-0.55***
		(0.14)
♯ of members over-report (Part I)		0.50**
		(0.26)
Constant	1.06***	2.17***
	(0.34)	(0.68)
n	69	69
Pseudo R^2	0.008	0.17
$\chi^2(1)$: *G-Charity* = *G-CharityR*	0.94	0.15

Notes. Instances where groups observed 5 points omitted.
***, ** and * denote $p < 0.01$, $p < 0.05$ and $p < 0.10$.

stop voicing arguments for dishonesty when the charity is involved.

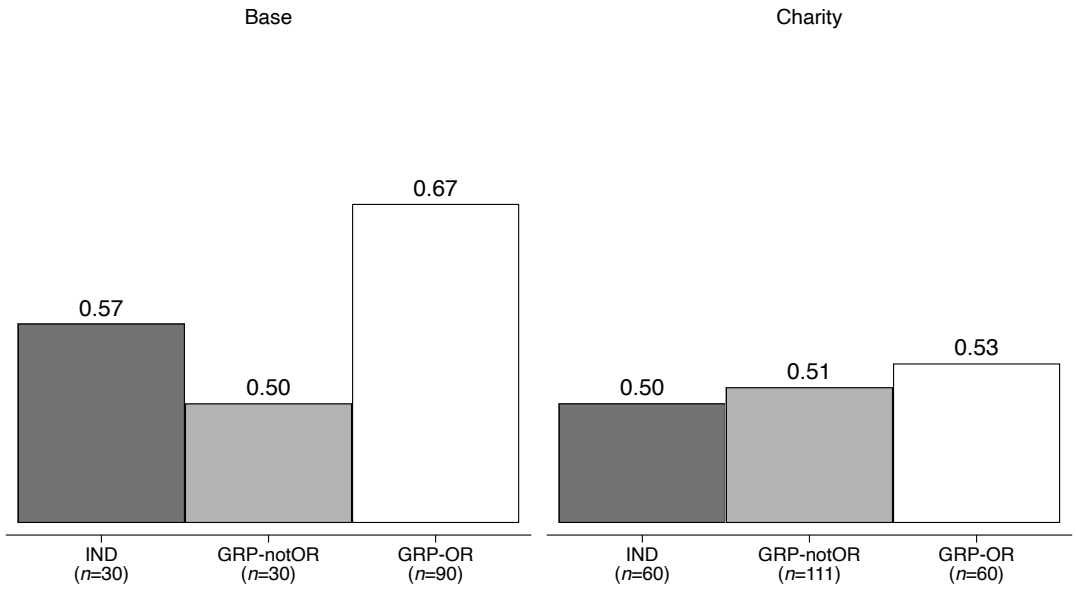
3.3 Spillovers of dishonesty

Finally, we study behaviour across parts to see whether dishonesty can spillover from one part to the next.

Do dishonest individuals make groups more dishonest? We first study whether over-reporting individuals make groups over-report more. We do not find any significant between-treatment differences in the number of group members who over-reported in Part I for groups in the *G-Base*, *G-Charity* and *G-CharityR* treatments (Kruskal-Wallis test, $p = 0.202$), so the above results cannot be explained by different group compositions.

Table 2 already showed us that there is a positive and significant relationship between the number of group members who over-reported in Part I and the proportion of statements for dishonesty in the chat.

This relation, however, does not seem to translate to actions: we find no significant correlation between the number of over-reporting subjects in Part I and over-reporting in Part II for groups in the *G-Base* (Spearman's $\rho = 0.235$, $p = 0.210$), *G-Charity* ($\rho = 0.117$, $p = 0.553$) and *G-CharityR* ($\rho = -0.117$, $p = 0.544$)



Notes. All subjects made individual decisions in Part III. IND: subjects who made individual decisions in Part II. GRP-notOR (resp. GRP-OR): subjects who made group decisions in Part II and whose group had not (resp. had) over-reported.

Figure 6: Proportion of IND, GRP-notOR and GRP-OR individuals who over-report in Part III.

treatments. The regression in Appendix B.2 confirms this result while further controlling for points observed in Part II.

We therefore have the following result:

Result 5. *The number of over-reporting individuals within the group has only a limited influence on the group's decision to over-report. The same is true even when the charity is involved.*

Do dishonest groups make individuals more dishonest? Finally, we look at whether subjects who were in a group that over-reported in Part II over-report themselves in Part III. For this we compare subjects who were in a group that over-reported in Part II (GRP-OR subjects) to subjects who were in a group that did not over-report in Part II (GRP-notOR subjects) and to subjects who were not in a group in Part II (IND subjects).

Figure 6 shows the proportion of IND, GRP-notOR and GRP-OR subjects who over-report in Part III. We pool the *Charity* and *CharityR* treatments together since subjects in these treatments faced the same situation in Part III. We see that the proportion of over-reporting GRP-OR subjects is only marginally higher than of IND and GRP-notOR subjects.

Table 3: Logit regression estimates: The spillovers from membership of dishonest groups.

Dependent Variable: Over-report in Part III.	BASE		CHARITY	
Regression	(1)	(2)	(3)	(4)
Reference group: IND subjects				
GRP-notOR	-0.66	-0.81	0.21	-0.45
	(0.58)	(0.66)	(0.36)	(0.49)
GRP-OR	1.19*	1.45***	1.01**	1.64**
	(0.65)	(0.70)	(0.51)	(0.53)
Points observed (Part III)		-0.41*		-0.75***
		(0.22)		(0.15)
Over-report (Part I)		-2.04**		2.97***
		(0.81)		(0.52)
Constant	0.88*	1.31*	0.51*	0.82
	(0.45)	(0.71)	(0.29)	(0.52)
n	96	96	186	186
Pseudo R^2	0.09	0.20	0.03	0.33
$\chi^2(1)$: GRP-DIS = GRP-notDIS	9.19***	10.08***	6.85***	5.01**

Notes. Instances where subjects observed 5 points are omitted.
Standard errors clustered at the group level.
***, ** and * denote $p < 0.01$, $p < 0.05$ and $p < 0.10$.

The regressions in Table 3 look at the impact of being in an over-reporting or a non-over-reporting group in Part II on the likelihood of over-reporting in Part III, controlling for behaviour in Part I, gender and points observed in Part III. The estimates show that GRP-OR subjects are significantly ($p \leq 0.068$) more likely to over-report in Part III relative to IND and GRP-notOR subjects in the *Base* and *Charity* treatments. Also, we do not find the likelihood of over-reporting to be significantly higher for GRP-notOR relative to IND subjects. This leads us to our final result:

Result 6. *Individuals are more likely to over-report if they were previously in a group that over-reported. This holds independently of whether the third party, such as a local charity, is harmed by the dishonest behaviour.*

4 Conclusion

We report the results of a laboratory experiment testing whether groups are more dishonest than individuals. We replicate the study by Kocher et al. (2018) but also add the fact that dishonest behaviour harms, not the experimental budget, but a

local charity. Our results show that once this incentive for honesty is introduced, groups are no more dishonest than individuals. In fact, groups can even help moderate the extent of dishonesty.

References

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond, "Preferences for truth-telling," *Econometrica*, 2019, *87* (4), 1115–1153.

Barr, Abigail and Georgia Michailidou, "Complicity without Connection or Communication," *Journal of Economic Behavior & Organization*, October 2017, *142*, 1–10.

Chytilová, Julie and Václav Korbel, "Individual and group cheating behavior: A field experiment with adolescents," Technical Report, IES Working Paper 2014.

Conrads, Julian, Bernd Irlenbusch, Rainer Michael Rilke, and Gari Walkowitz, "Lying and team incentives," *Journal of Economic Psychology*, 2013, *34*, 1–7.

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

_ **and Franziska Föllmi-Heusi**, "Lies in disguise – An experimental study on cheating," *Journal of the European Economic Association*, 2013, *11* (3), 525–547.

Gino, Francesca, Shahar Ayal, and Dan Ariely, "Self-serving altruism? The lure of unethical actions that benefit others," *Journal of Economic Behavior & Organization*, 2013, *93*, 285–292.

Gneezy, Uri, "Deception: The role of consequences," *American Economic Review*, March 2005, *95* (1), 384–394.

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

Kocher, Martin G., Simeon Schudy, and Lisa Spantig, "I lie? We lie! Why? Experimental evidence on a dishonesty shift in groups," *Management Science*, 2018, *64* (9), 3995–4008.

Mazar, Nina, On Amir, and Dan Ariely, "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance," *Journal of Marketing Research*, 2008, *45* (6), 633–644.

Muehlheusser, Gerd, Andreas Roider, and Niklas Wallmeier, "Gender differences in honesty: Groups versus individuals," *Economics Letters*, 2015, *128*, 25–29.

Papke, Leslie E. and Jeffrey M. Wooldridge, "Econometric methods for fractional response variables with an application to 401(k) plan participation rates," *Journal of Applied Econometrics*, 1996, *11* (6), 619–632.

_ **and** _ , "Panel data methods for fractional response variables with an application to test pass rates," *Journal of Econometrics*, 2008, *145*, 121–133.

Pruckner, Gerald J. and Rupert Sausgruber, "Honesty on the streets: A field study on newspaper purchasing," *Journal of the European Economic Association*, 2013, *11* (3), 661–679.

Schram, Arthur and Gary Charness, "Inducing social norms in laboratory allocation choices," *Management Science*, 2015, *61* (7), 1531–1546.

Soraperra, Ivan, Ori Weisel, Ro'i Zultan, Sys Kochavi, Margarita Leib, Hadar Shalev, and Shaul Shalvi, "The bad consequences of teamwork," *Economics Letters*, 2017, *160*, 12–15.

Weisel, Ori and Shaul Shalvi, "The collaborative roots of corruption," *Proceedings of the National Academy of Sciences*, 2015, *112* (34), 10651–10656.

Online Appendix

A Excess points

For the individual task, we define an individual's *excess points* as his received points minus his observed points. For the group task (i.e., part II of the *G-Base*, *G-Charity* and *G-CharityR* treatments), we define excess points as the points received points minus the observed points of a representative group member.[7]

Table A1 detail the mean excess point over Parts I, II and III of each treatment. Focusing on Part II, we find that excess points are significantly higher in the *G-Base* relative to *I-Base* treatments (Mann-Whitney $U = 341$, $p = 0.089$). In contrast, we do not find excess points for groups and individuals to be significantly different in Part II of the *Charity* (Mann-Whitney $U = 432$, $p = 0.774$) and *CharityR* (Mann-Whitney $U = 392$, $p = 0.376$).

The numbers reported in the main text convert these excess points to Euros by multiplying them by 2.

Table A1: Mean excess points.

	Base		*Charity*		*CharityR*	
	Ind.	Grp	Ind.	Grp	Ind.	Grp
Part I	1.20	0.98	1.57	1.13	1.13	0.67
	(2.09)	(1.70)	(1.96)	(1.93)	(1.94)	(1.63)
Part II	1.07	1.63	1.13	0.90	0.47	0.13
	(1.64)	(1.73)	(1.93)	(2.01)	(1.20)	(1.55)
Part III	1.67	1.77	1.63	1.43	0.93	0.90
	(1.73)	(1.75)	(1.59)	(1.76)	(1.87)	(1.96)

Notes. Each cell details the mean excess points with the standard errors in parenthesis. There are 30 observations in each cell of the Individual (*I-Base*, *I-Charity* and *I-CharityR*) treatments and part II of the Group (*G-Base*, *G-Charity* and *G-CharityR*) treatments—90 observations in each other cells of the group treatments.

[7]Subjects always receive the reported points in the individual task. In the group task, subjects only receive the reported points if all coordinated on the same report—otherwise 0 points.

B Additional regressions

In all regressions, we omit observations from the group treatments ($n = 6$ in *G-Charity* and $n = 3$ in *G-CharityR*) where group members fail to coordinate on reporting the same points in Part II.

B.1 Individuals vs. Groups

Regressions for Result 1. In Table B1, we use the Logit regression model to study subjects' likelihood of over-reporting in Part II controlling for points observed in Part II, gender and whether they had over-reported in Part I—standard errors clustered at matching group level.[8] The estimates find that subjects in the *Base* treatment are significantly ($p \leq 0.027$) more likely to over-report in Part II when they make decisions in a group. In contrast, subjects in the *Charity* ($p \geq 0.336$) and *CharityR* ($p \geq 0.592$) are NOT significantly more likely to over-report when they make decisions in a group.

We find no significant influence of gender ($p \geq 0.105$). Subjects are more likely to over-report in Part II if they had also done so in Part I. However, the effects are only significant in the *Base* treatments ($p = 0.047$) but not the *Charity* ($p = 0.367$) and *CharityR* ($p = 0.219$) treatments. There is a negative correlation between points observed in Part II and over-reporting in Part II. However, the correlation is only significant in the *Charity* ($p < 0.001$) and *CharityR* ($p = 0.004$) treatments and not the *Base* ($p = 0.873$) treatments. This discrepancy seems to be primarily driven by subjects in *G-Base* treatments whose decisions are insensitive to the observed points.[9]

Regressions for Result 2. Regressions (1) and (2) of Table B2 show that there are no significant differences in the likelihood of over-reporting between **i**ndividuals in the *I-Base* treatment and those in the *I-Charity* ($p \geq 0.753$), and between subjects in the *I-Base* treatment and those in the *I-CharityR* treatment ($p \geq 0.208$).[10] We also find no significant differences in the likelihood of over-reporting ($p \geq 0.338$) for **i**ndividuals in the *I-Charity* and *I-CharityR* treatments. Regressions (3) and (4) of Table B2 show that **g**roups are significantly more likely to over-report in the

[8]Subjects in the group treatments reported their decisions independently.

[9]We regressed over-reporting in Part II on points observed in Part II for subjects in the *G-Base* treatment. Whilst not significant ($p = 0.530$), the Logit model estimated coefficient is positive.

[10]We did not control for behaviour in Part I since subjects in the *Base* and *Charity* treatments face different dilemmas. Nevertheless, the conclusion will not change if we also controlled for such behaviour.

G-Base treatment relative to the *G-Charity* ($p \leq 0.040$) and *G-CharityR* ($p \leq 0.001$) treatments—no significant differences in the *G-Charity* and *G-CharityR* treatments ($p \geq 0.185$).

Regressions for Result 3. In Table B3, we use the Ordered Logit regression model to study the reported points (Part II) by over-reporting groups and individuals. We find no significant differences in the reported points of over-reporting **g**roups and **i**ndividuals in the *Base* ($p \geq 0.998$) and *CharityR* ($p \geq 0.853$) treatments. In contrast, we find reported points to be significantly lower ($p \leq 0.091$) for over-reporting **g**roups relative to **i**ndividuals in the *Charity* treatments.

B.2 Influence of dishonest members on group's decision.

In Table B4, we use the Logit regression model to study the likelihood of **g**roups over-reporting in Part II controlling for the points observed in Part II and the number of group members who over-report in Part I—we omit the proportion of statements for dishonesty as a covariate as it predicts the outcome perfectly. We find that the number of group members who over-reported in Part I has no significant ($p \geq 0.293$) influence on the **g**roup's likelihood over-reporting in Part II for all treatments.

Table B1: Logit model regression estimates: Over-reporting in Part II by subjects in the individual and group treatments.

Dependent Variable: Over-reporting in Part II						
	Base		Charity		CharityR	
Regression	(1)	(2)	(3)	(4)	(5)	(6)
Reference group: Subjects in the individual treatments.						
Group Treatment	1.61** (0.69)	1.63** (0.74)	0.35 (0.59)	0.82 (0.86)	0.06 (0.62)	0.40 (0.74)
Over-report (Part I)		1.36** (0.68)		0.58 (0.64)		0.75 (0.61)
Male		-0.01 (0.59)		1.13 (0.69)		0.17 (0.46)
Points Observed (Part II)		-0.03 (0.19)		-1.72*** (0.42)		-0.82*** (0.31)
Constant	-0.01 (0.42)	-0.24 (0.64)	-0.17 (0.41)	3.55*** (1.13)	-0.69 (0.43)	1.03 (0.87)
n	96	96	90	90	93	93
Clusters	48	48	46	46	47	47
Pseudo R^2	0.09	0.13	0.01	0.46	0.01	0.20

Notes. We exclude instances where 5 points were observed due to perfect collinearity. Standard errors are clustered at the matching group levels.
***, ** and * denote $p < 0.01$, $p < 0.05$ and $p < 0.10$, respectively.

Table B2: Logit model regression estimates: Over-reporting in Part II by individuals and groups.

Dependent Variable: Over-reporting in Part II				
	Individual treatments		Group treatments	
Regression	(1)	(2)	(3)	(4)
Reference group: Subjects in the *Base* treatment.				
Charity	-0.16 (0.57)	-0.19 (0.63)	-1.47** (0.69)	-1.69** (0.77)
CharityR	-0.69 (0.59)	-0.82 (0.65)	-2.23*** (0.70)	-2.57*** (0.79)
Points observed (Part II)		-0.61*** (0.19)		-0.78*** (0.24)
Constant	0.01 (0.40)	1.52** (0.68)	1.60*** (0.55)	3.94*** (1.01)
n	72	72	69	69
Pseudo R^2	0.02	0.13	0.13	0.27
$\chi^2(1)$: *Charity* = *CharityR*	0.78	0.92	1.75	1.58

Note. We exclude instances where 5 points were observed due to perfect colinearity.
***, ** and * denote $p < 0.01$, $p < 0.05$ and $p < 0.10$, respectively.

Table B3: Ordered Logit model regression estimates: Reported points in Part II by over-reporting individuals and groups.

Dependent Variable: Reported points in Part II (by individuals or groups)						
	Base		Charity		CharityR	
Regression	(1)	(2)	(3)	(4)	(5)	(6)
Reference group: Subjects in the individual treatments.						
Group Treatment	17.97	17.92	-2.03*	-2.03*	0.11	-0.18
	(5913.23)	(5868.03)	(1.19)	(1.20)	(0.93)	(0.97)
Points Observed (Part II)		0.12		-0.01		0.58
		(0.71)		(0.39)		(0.45)
n	32	32	23	23	16	16
Pseudo R^2	0.27	0.23	0.11	0.11	0.01	0.05

Notes. We exclude instances where 5 points were observed due to perfect collinearity.
***, ** and * denote $p < 0.01$, $p < 0.05$ and $p < 0.10$, respectively.

Table B4: Logit model regression estimates: Over-reporting in Part II by number of dishonest group members—group treatment only.

Dependent Variable: Over-reporting in Part II (Group treatment)			
	G-Base	*G-Charity*	*G-CharityR*
♯ of group members over-report in Part I	0.91	0.64	0.38
	(0.86)	(0.87)	(0.79)
Points observed (Part II)	0.09	-1.94**	-0.93**
	(0.39)	(0.76)	(0.41)
Constant	0.71	5.00**	1.51
	(1.13)	(2.36)	(1.09)
n	24	22	23
Pseudo R^2	0.07	0.54	0.22

Note. We exclude instances where 5 points were observed due to perfect colinearity.
***, ** and * denote $p < 0.01$, $p < 0.05$ and $p < 0.10$, respectively.

Table B1: Example 1. Extract of chat from a group in the *G-Charity* treatment.

Member	Points observed	Statement	x	time	label
1	4	number 4	4	9	Active
3	4	number 4	4	15	Active
2	4	number 4	4	23	Active
2	4	or highest payment number 5	5	38	Active
3	4	I would prefer number 4	4	53	Active
1	4	lets keep it 4	4	62	Passive
2	4	ok	4	67	Passive
3	4	is this okay for everyone?	4	74	Passive
2	4	yep	4	78	Passive
1	4	yes	4	79	Passive
3	4	ok than I will leave the chat :)			NA

C Analysis of chat data

The analysis of chat data was independently perform by a student assistant. The chat data was sorted by matching groups and time. The assistant followed the following steps.

- **Step 1.** Assign the label "active" to all statements where a subject makes a recommendation (i.e., points to report). Let $x \in \{0, 1, ., 5\}$ points be the recommendation made. If the statement refers to 2 or more possible recommendations, use the lowest recommedation.
- **Step 2.** Assign the label "passive" to all statements where a subject does not make a new recommendation but agrees with the previous recommendation. Let $x \in \{0, 1, ., 5\}$ be the most up to date recommendation points
- **Step 3.** Assign the label "NA" to all statements that are irrelevant to the reporting decision – we do not generate x for the NA statements.
- **Step 4.** Derive the proportion of statements for dishonesty, the frequency that x is less than the observed points for the matching group.

Tables B1 and B2 are two examples of the chat extracts. In the former, the proportion is $1/10 = 0.10$. In the latter, the proportion is $7/7 = 1$.

Table B2: Example 2. Extract of chat from a group in the *G-Base* treatment.

Member	Points observed	Statement	x	time	label
2	1	Hi		2	NA
3	1	Hi		7	NA
1	1	Hi		8	NA
2	1	You guys want to go for 5 points?	5	11	Active
1	1	Yes	5	22	Passive
3	1	ok	5	31	Passive
2	1	Alright its decided then	5	39	Passive
3	1	So we have to type 5 right?	5	44	Passive
2	1	Correct		49	NA
3	1	Ok		55	NA
3	1	So shall we leave the chat		63	NA
3	1			65	NA
1	1	Did you also have a 1 in the video?		68	NA
3	1	yes		74	NA
2	1	Yeah me too		82	NA
2	1	If you guys dont want to chat we can leave the chat		98	NA
3	1	So what exactly should we type in the box?		119	NA
2	1	Which box?		132	NA
1	1	I want to leave the chat open for the remaining time		132	NA
3	1	the Answer bo		143	NA
3	1	*box		151	NA
2	1	Type 5	5	152	Passive
3	1	ok	5	156	Passive
3	1	i would say we can leave the chat		182	NA
2	1	I guess Member 1 wanted to leave it open		216	NA
3	1	ok fine then		223	NA
2	1	Out of curiosity, what did you guys answer in Part 1?		229	NA
1	1	I just wanted to ask that		242	NA
1	1	I had a 1 in the video and I typed in 5		262	NA
3	1	me too		284	NA
2	1	I guess I was the only one with a 1 lol		298	NA

Table D1: Frequencies that points are observed.

	Part I		Part II		Part III	
	Our	KSS2018	Our	KSS2018	Our	KSS2018
0 points	30%	31%	10%	15%	10%	31%
1 point	10%	08%	10%	23%	20%	15%
2 points	10%	38%	20%	23%	10%	23%
3 points	20%	08%	10%	15%	30%	23%
4 points	10%	00%	30%	23%	10%	00%
5 points	20%	15%	20%	00%	20%	08%
$\chi^2(5)$	11.45^{**}		10.60^{**}		11.42^{**}	

Note. ***, ** and * denote $p < 0.01$, $p < 0.05$ and $p < 0.10$.

D Details on the experiment

D.1 Die-rolls

We pre-generated 10 sequences of die roll. Table D1 details the frequencies of points observed in our data. For completeness, we also report the frequencies of points observed in Kocher et al. (2018) data (KSS2018).

D.2 Instructions

The experiments were conducted in English. The "general procedures" were printed while the instructions for Parts I, II and III were displayed on the subjects' computer screens. Where relevant, the parts of the instructions that are unique to the Charity and CharityR treatments will be marked as "*text*". In addition, the instructions that are unique to CharityR treatments will be marked as text. Finally, we will refer to the charity as the XX charity.

D.2.1 Instructions: General procedures

Please switch off your electronic devices and remain silent. Also, do not talk with the other participants. For showing up on time you will receive a participation fee of 4 euros. You may also earn more during the experiment. The experiment consists of 3 parts (Part 1, Part 2 and Part 3). The three parts are independent: choices made in one part do not affect the other parts. At the beginning of each part you will see the detailed instructions for that part on your computer screen. If you have any questions, please raise your hand and an experimenter

will come to your desk to answer them. During the experiment, you and the other participants will make decisions. You may also interact with other participants, in which case your own decisions and the decisions of the others may determine your earnings. The onscreen instructions will clearly show whether you interact with other participants. They will also explain how exactly your earnings will be determined.

Payment. In some part of the experiment, and depending on your decisions, you will earn points for yourself *or for a charity. We will provide further information about this charity in a minute.* At the end of the experiment, only the points from one of the 3 parts will be used to determine the payment to you *and to the charity*. Your points *and the charitys points* in this part will be converted into euros at the exchange rate of:

1 point = 2 Euros

To select the part for payment, the computer will randomly ask one participant to roll a die:

- If the die shows a ⚀ or ⚁, then points from Part 1 will be used for payment;
- If the die shows a ⚁ or ⚄, then points from Part 2 will be used for payment;
- If the die shows a ⚂ or ⚅, then points from Part 3 will be used for payment.

After converting points into euros, we will pay you your total earningsyour earnings from the selected part and the participation fee of 4 Euro. No other participant will learn about your earnings and you will not learn about the earnings of others.

We will also pay the charity via online transfer. We will do so from the experimenter room, and you are welcome at this stage to come monitor the payment and verify how much we donate.

Anonymity. We will never link your name with the decisions you will make in this experiment. You will not learn the identity of the other participants, and the other participants will not learn your identity. At the end of the experiment we will ask you to sign a receipt to confirm the payments you received and the payments for the charity that are determined by your decisions. We only use this receipt for accounting and it is not linked to your decisions.

The XX charity *Your decisions during the experiment will affect XX, a local charity based in Nuremberg. In the next few lines, we wish to give you more information about this charity and its goals.*

XX is a group of clowns that travel to the hospitals of the Franconian region to visit sick children and brighten their day. They visit children who have been hospitalised for a short time, as well as children who are seriously or chronically ill and in intensive care, oncology, cardiology or dialysis.

The clowns visit the hospitals-Klinikum Nuremberg-Süd, Klinikum Fuerth, and the University Hospital in Erlangen—at least once every two weeks and sometimes every week. During their visit they do not perform a rehearsed program but instead interact spontaneously with each child in their room. They are also in close contact with physicians, nurses, educators and psychologists, in order to adapt their visit to the needs of every children.

At the moment there are 9 clowns in the charity and they are all volunteers. 100% of the donations they receive go directly to their work as clowns in the hospitals. For example, the donations pay for the red noses, the makeup, the transportation costs to the hospitals, the flyers, and the website. XX was founded in 1999 and is recognised by the tax office of Nuremberg as a non-profit organisation particularly worthy of promotion.

As a result of your decisions XX will receive some points. How many exactly will be detailed in the instructions that will appear on your computer screen at the beginning of each Part. At the end of the experiment and before paying you your earnings we will add all the points received by XX and convert them to euros. As explained above we will donate this amount to XX via online transfer.

Before we start the experiment, you will see appear on your computer screen the webpage of XX. We will give you 5 minutes to check their website. It is only in German but you can see the pictures of the clowns in "Wir Clowns" and of their visits in "Clown Nachrichten". At the top of the page you will also find links to their Facebook and Youtube, where you can see more news, pictures, and videos.

D.2.2 Instructions: Part I

In Part 1 of the Experiment, you will see a video of a die roll showing ⚀, ⚁, ⚂, ⚃, ⚄ or ⚅. The video you will see is randomly selected by the computer and each of the 6 possible videos is chosen with equal probability.

Your task. Your task is to remember the number of the die roll in the video and to type it into a field showing up later. The die number you enter determines YOUR points *and the points for the CHARITY as explained by the table below* (see

	⚀	⚁	⚂	⚃	⚄	⚅
Die number entered	1	2	3	4	5	6
Points for you	1 point	2 points	3 points	4 points	5 points	0 points
Points for the charity	*4 points*	*3 points*	*2 points*	*1 point*	*0 points*	*5 points*

Table D2: Points Table in Parts 1, 2 and 3.

Table D2). For example, if you enter the die number to be "4", you will receive 4 points *and the charity will receive 1 point.*

Comprehension questions. **remark: Subjects had to correct answers the questions to begin part 1. The correct answer is underlined.**

1. What is your task in this part?
   - To enter the number that you have seen and memorised.
   - To enter a number different from the number you have seen and memorised.
   - To enter an arbitrary number.

2a. Suppose that you see a ⚂ and enter a "3". How many points will YOU receive? 3 points.

2b. *How many points will the CHARITY receive?* 2 points.

3a. Suppose that you see a ⚂ and enter a "2". How many points will YOU receive? 2 points.

3b. *How many points will the CHARITY receive?* 3 points.

4a. Suppose that you see a ⚂ and enter a "4". How many points will YOU receive? 4 points.

4b. *How many points will the CHARITY receive?* 1 point.

D.2.3 Instructions: Part II (individual treatments)

Part 2 of the experiment is similar to Part 1, except that now you have the opportunity to record your thoughts BEFORE making your decision.

In Part 2 of the Experiment, you will see a video of a die roll showing ⚀, ⚁, ⚂, ⚃, ⚄ or ⚅. The video you will see is randomly selected by the computer and each of the 6 possible videos is chosen with equal probability.

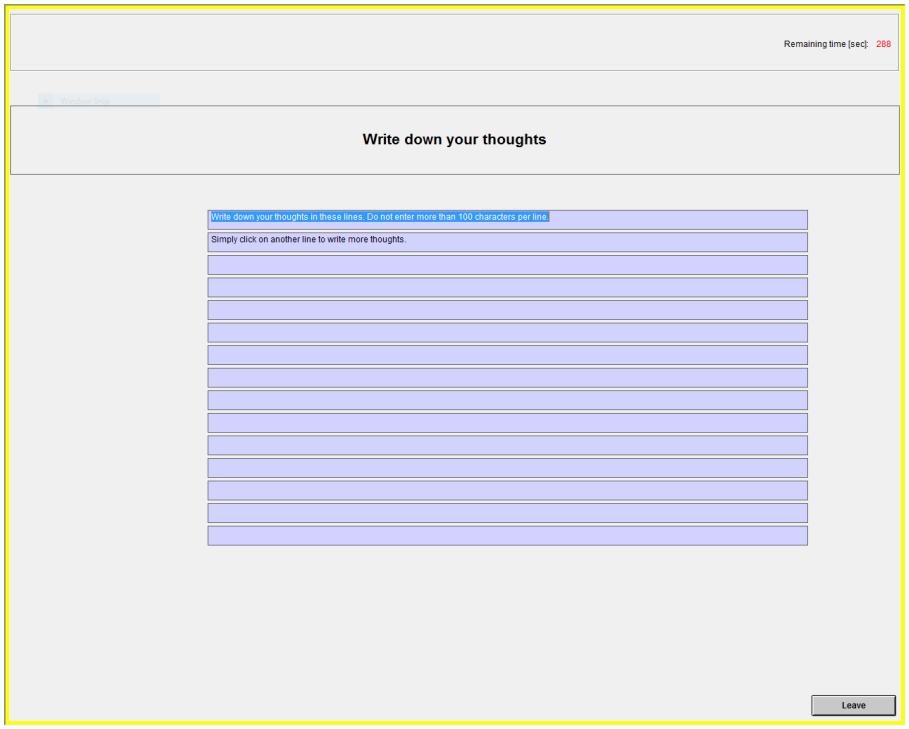


Figure D1: How subjects entered their thoughts.

Your task. Your task is to remember the number of the die roll in the video and to type it into a field showing up later. The die number you enter determines YOUR points *and the points for the CHARITY as explained by the table below* (see Table D2). For example, if you enter the die number to be "4", you will receive 4 points *and the charity will receive 1 point.* The next screen will describe how you can record your thoughts.

Comprehension questions. **remark: Subjects had to correct answers the questions. The questions are exactly the same as in Part I.**

How you input your thoughts. You have the possibility to record your thoughts about the number you will enter. You have 5 minutes to write down your thoughts. After 5 minutes the possibility will end. If you have finished before the 5 minutes are over, you can click on the "Leave" button (subjects see Figure D1).

After the experiment we will send a copy of what you wrote to the charity. We are distributing an illustration of the copy for you to check what kind of information we will send to the charity. As you will see, we will send the number you saw, the number you reported, and what you wrote. Note that, since the experiment is anonymous, the copy is anonymous as well: only the participant ID appears and it cannot be traced to you. We will send the transcript to the charity via email

and we will add the email addresses of everyone who participated to today's experiment in blind carbon copy (so they will not see your email address) for you to verify that we are really sending the copy.

D.2.4 Instructions: Part II (group treatments)

Part 2 of the experiment is similar to Part 1, except that now you decide in a group. We will randomly match you with 2 other participants such that you form a group of 3.

In Part 2 of the Experiment, you will see a video of a die roll showing ⚀, ⚁, ⚂, ⚃, ⚄ or ⚅. The video you will see is randomly selected by the computer and each of the 6 possible videos is chosen with equal probability. ALL MEMBERS OF YOUR GROUP WILL SEE THE SAME RANDOMLY CHOSEN VIDEO.

Your task. Your task is to remember the number of the die roll in the video and to type it into a field showing up later. YOU *and the CHARITY* will receive points from this task only when all group members enter the same number. In contrast, YOU *and the CHARITY* will receive 0 points if any group member enters a different number. If all group members enter the same number, the number entered determines the points for YOU *and the CHARITY* as described by the table below (Table D2). For example, if all group members enter the number 4, then each group member will receive 4 points. *In addition, the CHARITY will receive 1 point from each group member (the charity receives a total of 3x1=3 points).* If any group member enters a different number, each group member receives 0 points *and the charity also receives 0 points from each group member (the charity receives a total of 3x0=0 points).* You will be able communicate with the other group members. How you do so will be explained on the next screen.

Comprehension questions. **remark: Subjects had to correct answers the questions to proceed. The correct answer is underlined.**

1. What is your task in this part?
   - To enter the number that you have seen and memorised.
   - To enter a number different from the number you have seen and memorised.
   - To enter an arbitrary number.
2. Suppose that you see a ⚃ and enter a "3". The others also enter "3".

(a) Points YOU receive. 3 points.

(b) *Points the Charity receives from YOU* 2 points.

(c) *Points the Charity receives from your GROUP* 6 points.

3. Suppose that you see a ⚂ and enter a "2". The others also enter "3".

(a) Points YOU receive. 2 points.

(b) *Points the Charity receives from YOU* 3 points.

(c) *Points the Charity receives from your GROUP* 9 points.

4. Suppose that you see a ⚃ and enter a "4". Someone enters "5".

(a) Points YOU receive. 0 points.

(b) *Points the Charity receives from YOU* 0 points.

(c) *Points the Charity receives from your GROUP* 0 points.

Group interaction. You have the possibility to communicate with the other two group members via a chat box to clarify the number each group member will enter. You have 5 minutes to communicate. The group discussion ends after the 5 minutes or as soon as all 3 members of the group have pressed the "leave chat" button. If only 1 or 2 members of the group press the button, the discussion will continue. The group discussion will only end if all members press the button or if time runs out. If you have pressed the button "leave chat" but you do not want to leave the chat, you can press the button "back". After the group discussion, each member of the group enters a number on the screen.

Generally, the course of communication is up to you. You may chat in any language as long as all group members understand the language. However, you are not allowed to make threats or to agree upon side payments within your group. If you are breaking these rules, you will be excluded from the experiment and you will not receive any payment from the entire experiment.

Within the given time, you can send as many messages to the other group members as you like. The messages you send appear automatically on the screens of your other group members. You cannot send a message to one member in particular.

The screen of the chat will look like this: (subjects see Figure D2) To write a message, click on the purple field, enter your message and press "Enter". Then, your message appears in the grey field above the purple field. You can send as

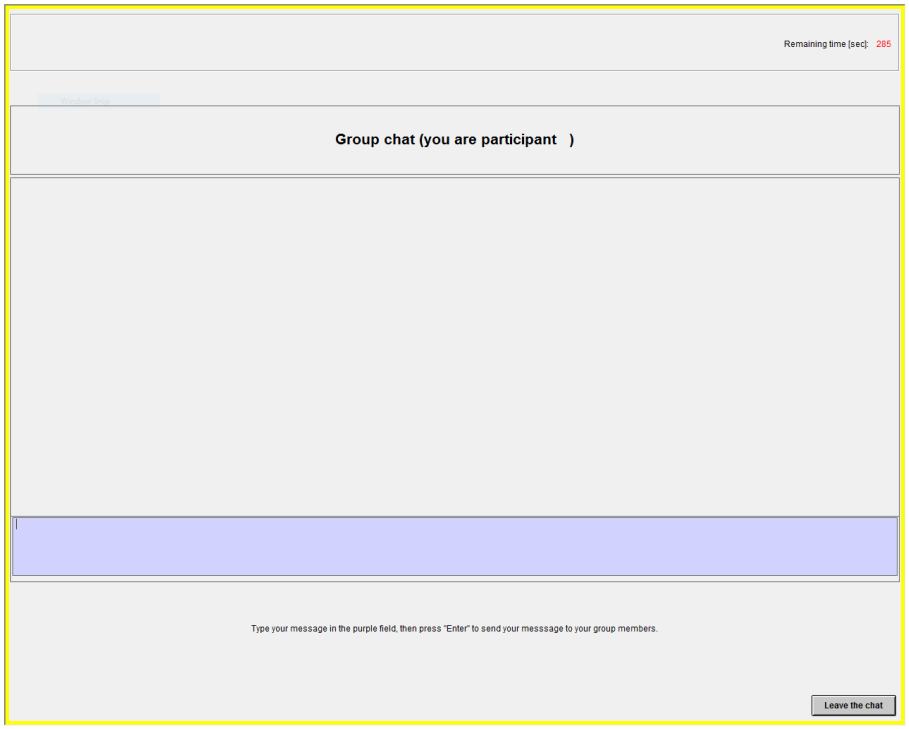


Figure D2: How subjects chat.

many messages as you want using the same procedure. The other participants will see your message only when you have pressed "Enter".

After the experiment we will send a copy of the chat to the charity. We are distributing an illustration of the copy for you to check what kind of information we will send to the charity. As you will see we will send the number your group saw, the number you reported, and the chat messages you sent. Note that, since the experiment and the chat are anonymous, the copy is anonymous as well: only the participant and group ID appear and these cannot be traced to you. We will send the transcript to the charity via email and we will add the email addresses of everyone who participated to today's experiment in blind carbon copy (so they will not see your email address) for you to verify that we are really sending the copy.

D.2.5 Instructions: Part III

Part 3 of the experiment is the same as Part 1. That is, your task in Part 3 is exactly the same as in Part 1.

In Part 3 of the Experiment, you will see a video of a die roll showing ⚀, ⚁, ⚂, ⚃, ⚄ or ⚅. The video you will see is randomly selected by the computer and each of the 6 possible videos is chosen with equal probability.

Your task. Your task is to remember the number of the die roll in the video and to type it into a field showing up later. The die number you enter determines YOUR points *and the points for the CHARITY as explained by the table below* (see Table D2). For example, if you enter the die number to be "4", you will receive 4 points *and the charity will receive 1 point.*

Comprehension questions. [Note: Subjects had to correctly answers the questions. The questions are exactly the same as in Part I.]