Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

I Lie? We Lie! Why? Experimental Evidence on a Dishonesty Shift in Groups

http://orcid.org/0000-0002-6792-1192Martin G. Kocher, http://orcid.org/0000-0002-4409-0576Simeon Schudy, http://orcid.org/0000-0003-0776-3863Lisa Spantig

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–14
ISSN 0025-1909 (print), ISSN 1526-5501 (online)
http://pubsonline.informs.org/journal/mnsc/

I Lie? We Lie! Why? Experimental Evidence on a Dishonesty Shift in Groups

Martin G. Kocher,[a,b,c] Simeon Schudy,[b] Lisa Spantig[b]

[a] Institute for Advanced Studies, 1080 Vienna, Austria; [b] Department of Economics, Ludwig-Maximilians-University of Munich, 80539 Munich, Germany; [c] University of Gothenburg, SE-405 30 Gothenburg, Sweden

Contact: martin.kocher@econ.lmu.de, http://orcid.org/0000-0002-6792-1192 (MGK); simeon.schudy@econ.lmu.de, http://orcid.org/0000-0002-4409-0576 (SS); lisa.spantig@econ.lmu.de, http://orcid.org/0000-0003-0776-3863 (LS)

Received: August 1, 2016
Revised: January 24, 2017
Accepted: February 28, 2017
Published Online in Articles in Advance: August 3, 2017

https://doi.org/10.1287/mnsc.2017.2800

Abstract. Unethical behavior such as dishonesty, cheating and corruption occurs frequently in organizations or groups. Recent experimental evidence suggests that there is a stronger inclination to behave immorally in groups than individually. We ask if this is the case, and if so, why. Using a parsimonious laboratory setup, we study how individual behavior changes when deciding as a group member. We observe a strong dishonesty shift. This shift is mainly driven by communication within groups and turns out to be independent of whether group members face payoff commonality or not (i.e., whether other group members benefit from one's lie). Group members come up with and exchange more arguments for being dishonest than for complying with the norm of honesty. Thereby, group membership shifts the perception of the validity of the honesty norm and of its distribution in the population.

History: Accepted by Uri Gneezy, behavioral economics.
Funding: L. Spantig acknowledges funding through the International Doctoral Program "Evidence-Based Economics" of the Elite Network of Bavaria.
Supplemental Material: Data and the online appendix are available at https://doi.org/10.1287/mnsc.2017.2800.

Keywords: dishonesty • lying • group decisions • communication • norms • experiment

"I did steal from Enron. We *stole from Enron."*
—The Enron Trial: Testimony of Andrew Fastow (former CFO of Enron)[1]

"The conduct was fairly open and notorious, I would say. It was no great secret what we *were doing."*
—Christopher Loehr (former analyst for Enron)[2]

1. Introduction

Groups and organizations sometimes fail to comply with a moral norm. They lie, they cheat, they are dishonest, they are corrupt, and they commit fraud. However, it is not organizations that take those decisions; it is individuals that are part of the organization. Can we thus explain undesired behavior in organizations simply by aggregating individual failures to comply with the norm? Or, are there other elements inherent to the organization or to its structure that can help us better understand how undesired behavior emerges? And, how does undesired behavior of individuals differ from behavior within or by an entire organization? While these are relevant questions, surprisingly little empirical evidence exists (Conrads et al. 2013, Sutter 2009). This paper addresses these questions in a parsimonious setup that allows us to identify some of the potential reasons for collective failure to follow a moral norm or to comply with desired behavior.

Recent years have provided several prominent examples of unethical behaviors in groups and organizations. Fraudulent accounting methods and malpractice of groups of executive officers have led to the marked bankruptcies of WorldCom and Enron. More recently, it has been discovered that *inter alia* the German car producer Volkswagen has sold diesel cars with emissions certificates based on potentially faulty information (Wallace 2015). To improve emission test results, Volkswagen has allegedly installed software in their diesel engines that could detect when the cars were on the test stand and adjust the engine performance accordingly. After investigations by the Environmental Protection Agency (EPA), newspaper articles say that, "VW must have had a chain of management command that approved fitting cheating devices to its engines" (see, e.g., Hotten 2015). However, it is not only for-profit firms that are involved in unethical behavior: there are cases of charities that commit embezzlement, sports organizations and executives that generate financial scandals or engage in morally and legally questionable practices in the context of doping, and sports teams that violate established norms (see, e.g., *BBC* 2016, *Economist* 2015, Ruiz 2015).

This study provides a twofold contribution. First, we implement a parsimonious laboratory setup to investigate whether groups (as our proxy for small organizations) are indeed more inclined to engage in dishonest or unethical behaviors than individuals, as casual observation and some previous results in the literature suggest (e.g., Chytilova and Korbel 2014, Conrads et al. 2013, Gino et al. 2013, Muehlheusser et al. 2015, Sutter 2009, Weisel and Shalvi 2015). We find that the answer is affirmative. Individuals lie less frequently when deciding alone as compared to groups. Second, we discuss and single out explanations for the "dishonesty shift" in groups. There are several candidate explanations: (i) a simple aggregation of individual inclinations as a consequence of aggregation rules (i.e., decision-making procedures) within the group; (ii) the incentive structure inherent to many group decisions (oftentimes, all members share group payoffs equally and an individual deviation from either of the strategies—behaving dishonestly or honestly—can sometimes reduce payoffs for everyone dramatically); (iii) the decreased observability of one's actions within a group, potentially making the individual less accountable for their actions when they are group members; and (iv) the deliberation process inherent to group interaction involving, for instance, the exchange of arguments and learning about the strength and prevalence of a norm.

Our laboratory experiment uses a variant of the die-rolling task introduced by Fischbacher and Föllmi-Heusi (2013). Participants are asked to report the result of a die roll, and their payoffs depends on their reports.[3] Hence, participants face a trade-off between being honest by reporting the true number and potentially forgoing a monetary profit (i.e., following the norm of honesty), on the one hand, and being dishonest and potentially earning more (i.e., violating the norm of honesty), on the other hand. We implement individual decision-making situations and group decision-making situations using an experimental design that allows us to study behavioral change (within subjects) across several (between subjects) treatments. Our setup reduces the effects from the decreased observability of one's actions as a member of a group and enables us to assess whether the exchange of arguments and learning about the strength and prevalence of a norm or whether incentive structures and preference aggregation are the main drivers of the observed dishonesty shift.

In the die-roll paradigm, a pure payoff maximizer would want to always report the number that yields the highest monetary payoff, regardless of the actual die roll. This is true for both the individual and the group decision-making situation. If one assumes sufficiently high moral costs of lying, individuals might want to report truthfully, and the group outcome then depends on the aggregation of individual preferences. For instance, unanimity should lead to less lying than other aggregation mechanisms. Adding social image concerns, accountability considerations and changes in the perception of the norm, group interaction can drag the comparison between the individual and the group setting in any direction—toward more or less dishonest behavior (Bénabou 2013; Bénabou and Tirole 2006, 2012; Dufwenberg and Dufwenberg 2016; Falk and Tirole 2016). Ultimately, the question of whether groups or individuals are more or less dishonest (and the causes and consequences of any potential shift from the individual to the group setting) can only be answered empirically.

A laboratory experiment has several advantages when it comes to the identification of the effects we are interested in: it allows us (i) to exogenously vary group membership and incentives for the group members to behave dishonestly, (ii) to observe individual behavior before individuals become members of a group as well as behavior/communication when the group decides, and (iii) to elicit individual beliefs about dishonest behavior of others. Naturally, there are limits to laboratory experiments. The task that we use is specific in the sense that it implies a zero fine for violating the norm of honesty. In bearing with the growing experimental literature on cheating and lying, this enables us to compare our results to existing studies that focus on individual decisions to lie. We implement group decisions in small groups with anonymous real-time chat interactions to keep as much experimental control as possible. These design choices lend themselves to extensions that bring the experimental setup closer to existing organizations with their hierarchies and with face-to-face interactions. Ultimately, field experiments are a desired methodology. Our aim here is to establish a set of explanations in a rigorously controlled environment that feeds into the design of future studies, relaxing some of our restrictions systematically.

The results from our setup reinforce the conclusion from the small existing literature: groups are (much) more inclined to lie than individuals. Drawing on the older psychology literature that detected a shift from individuals to groups in terms of risky decision making, coining it the "risky shift" (Pruitt and Teger 1969, Teger and Pruitt 1967), we refer to our finding as the "dishonesty shift." However, our main contribution is in providing an explanation for this shift: the shift can be explained neither by a higher level of strategic sophistication of groups than individuals (Sutter 2009), nor by a decreased observability of one's action as a group member. Both explanations are excluded by our design. Using appropriate treatment variations, we can also rule out that groups lie more than individuals because the other group members benefit from lying (Gino et al. 2013) or because group decisions require unanimity.

Our findings provide strong evidence for the importance of two mechanisms that drive the dishonesty shift and that are inherent to almost any group interaction: communication and learning about norm compliance. Communication exposes group members frequently to arguments in favor of violating the norm. The exchange of arguments and talking to people that argue in favor of violating the norm also changes the norm perception. We show that the expectation that other people (out-of-sample) lie increases significantly after the group interaction. A detailed analysis of the protocols from the group interaction suggests that groups lie more because communication enables them to justify dishonest behavior in a different way than individuals. Further, we find that the dishonesty shift in groups is very strong such that the group composition (in terms of the number of initially dishonest group members) only weakly affects the extent of dishonesty in a group.

Understanding the mechanisms that contribute to the dishonesty shift is essential, as it is a prerequisite for designing institutions and incentives that are conducive to norm compliance. Our results show that the availability and exchange of arguments that justify norm-violating behavior is an important aspect. Such exchange is occurring naturally in groups where group members discuss how they should act. A next step could be to analyze potential mechanisms in the group interaction that are able to counterbalance this effect, such as reminders of the norm or other related interventions.

The remainder of the paper is organized as follows: Section 2 describes the details of our experimental design and procedures, gives an overview of the literature, and provides behavioral predictions; Section 3 presents the results from our experiment; Section 4 discusses our findings; and Section 5 concludes.

2. Experimental Design, Related Literature, and Predictions

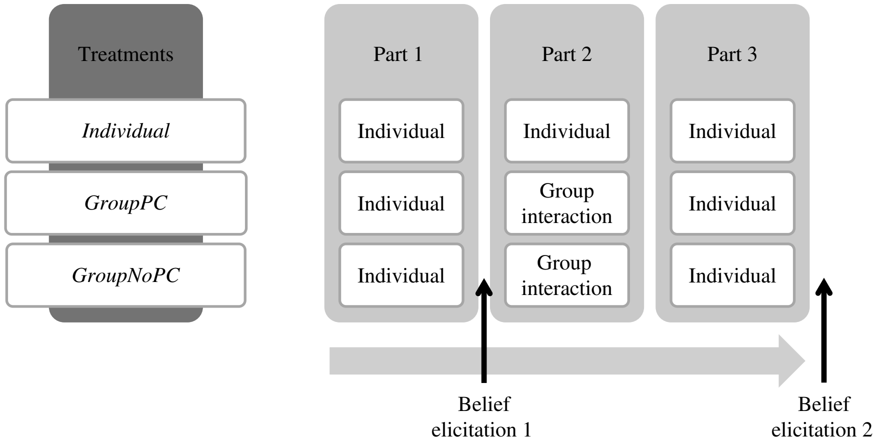
2.1. Experimental Design and Procedures

We use a variant of the die-rolling task introduced by Fischbacher and Föllmi-Heusi (2013), in which a pure payoff maximizer would want to always report the number that yields the highest monetary payoff (irrespective of the observed die roll). Experimental participants see a video of a die roll on their computer screen. The computer randomly chooses one out of six videos, each showing one possible outcome. Participants are informed about this—i.e., they know that the computer chooses each video with the same probability. The outcome of the die roll in the video is clearly visible for about 10 seconds. The participants' task is to enter the outcome shown in the video on the next screen, in a field stating "die number seen: __." Participants can enter any number between 1 and 6. Importantly, payoffs depend on the number entered, not on the number actually seen. We use the same payoff structure as the original experiment—i.e., die numbers ⚀ to ⚄ yield one to five points, respectively, and die number ⚅ yields zero points. As participants can enter any number between one and six, they have the possibility to report dishonestly. As we explicitly asked participants to report the number *seen*, we refer to deviating from truthful reporting as "dishonesty" or "lying" for the rest of the paper.

Die rolls shown in the video are chosen randomly by the computer and known by the experimenter such that participants cannot disguise their lies.[4] Thus, in contrast to the original die-rolling task, misreporting the die number in our experiment is clearly dishonest, but liars cannot disguise (for a similar task with observability, see also Gneezy et al. 2016). This is a desirable feature for our purposes for at least two reasons. First, the paradigm ensures that groups do not lie more than individuals because it is *easier to disguise* lies in a group, as individual decisions are clearly observable by the experimenter in individual and group treatments. Second, the paradigm allows us to study in a within-subject design how individual behavior changes—i.e., when reporting alone as compared to when reporting after communication in a group. Further, the task is easy to understand, and we can thus exclude that groups have a significantly *better understanding* of the task. Nonetheless, full observability may affect the level of lying (as compared to no observability). Our main analysis focuses, therefore, on the relative comparison between individual decision making and group decision making (and not on the absolute levels of dishonest behavior). Interestingly, direct comparisons between situations of full observability by the experimenter and full privacy (Gneezy et al. 2016) show minor differences in the full extent of lying, and similar conclusions can be drawn from comparable experiments that introduce an anonymous observer of the private die roll other than the experimenter (see, e.g., Baeker and Mechtel 2015, Houser et al. 2016, Van de Ven and Villeval 2015).

We implement a mixture of a within-subject and a between-subject design with three different between-subject treatments—*Individual*, *GroupPC*, and *Group NoPC*. Figure 1 illustrates the experimental design. Each treatment consists of three independent parts, for which instructions are displayed on screen at the beginning of each part. One of the three parts is randomly determined to be payoff relevant at the end of the experiment, and this is common knowledge. In the *Individual* treatment, participants make the same decision in each of the three parts: they see a video of a die roll on their screen and are asked to enter the die

Figure 1. Experimental Design and Elicitation of Beliefs



Note. Subjects are randomly allocated to treatment *Individual*, *GroupPC*, or *GroupNoPC* (between-subject) and participate in all three parts (within-subject).

number seen on the next screen. In the two group treatments, the first part and the third part are equivalent to the decision in the same parts in *Individual*. In the second part of the group treatments, subjects decide in groups of three participants that are randomly assembled. Each participant in a group sees the same video (and this is common knowledge). After group members have seen the video, they have five minutes' time to discuss in a group chat. The real-time chat allows for free-form communication without revealing one's identity (excluding the possibility of threats and side payments). All three group members see all messages that are sent in the chat. The three members can also decide unanimously to leave the chat before its automated ending. After the chat, each participant privately enters the number (as in *Individual*), making each individual decision fully observable by the experimenter in all treatments. The two group treatments reflect different organizational structures as we vary whether or not group members earn a common payoff. In *GroupPC*, participants face a decision with payoff commonality. Each member of a group has to enter the same number to receive a payoff, and each participant receives the payoff that corresponds to the number entered. If entered numbers differ within a group, all group members receive zero payoffs. Such a protocol implements a strong unanimity component. In *GroupNoPC*, participants face no payoff commonality—i.e., they receive payoffs according to the number they enter, irrespective of the number entered by the other group members. *GroupNoPC* is thus identical to *Individual* except for the group chat after having seen the same video. It is important to notice that there is no pressure to contribute to the chat and that the content of the chat is totally up to the participants (i.e., they could talk about the weather, the weekend, and anything they wanted, except for revealing their identities in any way).

As the first part is identical in all treatments (and also decisions are very similar), we can use the between-subject comparisons between *Individual* and *GroupPC* as well as *Individual* and *GroupNoPC* in part 2 to establish the difference in dishonest behavior between individual and groups (i.e., the potential existence of the dishonesty shift). The comparison between behavior in *GroupPC* and *GroupNoPC* in part 2 allows us to address the effect of payoff commonality—i.e., the relevance of the other-regarding concerns argument in group decision making that involves a trade-off between payoff maximization and norm compliance.[5]

To observe whether the group chat changes participants' beliefs about others' behavior in terms of (dis)honesty, we elicit our participants' beliefs about the lying behavior in a past experiment—i.e., in the baseline condition of Fischbacher and Föllmi-Heusi 2013. Following Fischbacher and Föllmi-Heusi (2013), we elicit participants' beliefs about the distribution of payoffs. We inform our participants that they have to guess the behavior of other participants in a similar, previously run experiment (from now on "reference experiment"). Participants then guess the shares of participants that earned a specific payoff—i.e., zero points, one point, two points, etc. Implicitly, the distribution provides the belief about honesty of similar decision makers. Participants earn five euros if they guessed all shares correctly. For every percentage point deviation from the correct shares, we reduce participants' payoff by 0.04 euros. The minimum payoff they can earn in the belief elicitation task is 0.50 euros. To avoid potential problems of hedging, we randomly selected one of the two belief-elicitation tasks at the end of the experiment to be payoff relevant.[6] Two aspects of the belief elicitation are important: first, the way we elicit beliefs makes it clear to participants that we ask them about past behavior of other participants, which is exogenous to the current experiment; second, we make clear that the data from the reference experiment is exactly the same for belief elicitation 1 and belief elicitation 2. Hence, changes in elicited values indicate a change in beliefs that can only be explained by the experience in our experiment.

During our experiment, each participant sat at a randomly assigned, separated PC terminal and received a copy of printed general instructions up front.[7] These informed participants that the experiment consists of three independent parts (one randomly determined to be payoff relevant at the end) and that the specific instructions for each part would be displayed at the beginning of each part on the participants' computer screen. A set of on-screen control questions ensured the understanding of the game in each part.[8] Each part started only after all subjects answered the control questions correctly. Participants could fail the control questions repeatedly and were allowed to ask the experimenter to provide an explanation. No form of communication was allowed during the experiment (except through the computerized chat environment in the group treatments). We conducted all sessions at the Munich Experimental Laboratory for Economic and Social Sciences (MELESSA) at Ludwig-Maximilians-University of Munich. The MELESSA subject pool includes undergraduate and graduate students of all fields of study. The data for the main treatments was collected over 14 sessions between June and September 2015, with 273 participants in total (39 in *Individual*, 117 in *GroupPC*, and 117 in *GroupNoPC*).[9] In the first two sessions, the computer randomly selected for each group in each part one out of the six possible videos of a die roll with equal probability. To increase statistical power and simplify nonparametric comparisons between treatments, we used these randomly determined sets of videos in the later sessions, such that in each part, the same die rolls were observed in each treatment. This procedure ensures that die rolls displayed are held constant across treatments, and we can compare the number of dishonest reports across treatments, holding the monetary costs of being honest constant. In our main analysis, we compare the individual-level data of the control treatment to the collapsed data from the group treatments.[10] Hence, we have 39 statistically independent observations in each treatment.

Participants received a show-up fee of four euros that was added to the earnings from the experiment. Subjects could earn points, where one point was equal to two euros. The experiment took about an hour. The average income (including the show-up fee) amounts to 14.62 euros. The experiment was programmed and conducted using z-Tree (Fischbacher 2007). We recruited participants using the online recruiting system ORSEE (Greiner 2015) and excluded all subjects with previous experience in die-rolling tasks or other similar experiments.

2.2. Related Literature and Behavioral Predictions

In our task, a rational selfish individual would want to report a "⚄," regardless of the actual die roll. This is true for both the individual and group decision-making situations. The experimental literature in economics has shown that frequently individuals are willing to forego monetary benefits to behave honestly in such situations (see, e.g., Abeler et al. 2014; Cappelen et al. 2013; Erat and Gneezy 2012; Fischbacher and Föllmi-Heusi 2013; Glätzle-Rützler and Lergetporer 2015; Gneezy 2005; Gneezy et al. 2013, 2016; Kröll and Rustagi 2016; Lundquist et al. 2009; Mazar et al. 2008). This contrasts behavior of purely self-interested payoff maximizers but is in line with models that incorporate moral costs of lying.

Recent work has started to investigate dishonesty in groups (e.g., Baeker and Mechtel 2015, Chytilova and Korbel 2014, Conrads et al. 2013, Muehlheusser et al. 2015, Sutter 2009). The group decision-making setup adds several dimensions to the problem. First, it adds the aggregation problem of individual preferences in case they are not completely aligned. Second, the group setup might make social aspects more relevant: hiding behind the other group members (accountability), payoff commonality and the need to coordinate, and social imagine concerns. Based on existing work, we may expect groups to lie more than individuals for at least three reasons. First, groups may generally apply significantly higher levels of reasoning than individuals (Kocher et al. 2006, Kocher and Sutter 2005). Thus, groups may lie more as they have a *better understanding* of the game (Sutter 2009). Second, groups may lie more, as it can be *easier to disguise* lying in groups than individual lies (Conrads et al. 2013). Both arguments should play a minor role in the context of our experimental design. The task is easy, and individual choices are perfectly observable by the experimenter in all treatments. Third, recent work suggests that groups could lie more because *others may benefit* as well from dishonest behavior (Gino et al. 2013, Weisel and Shalvi 2015, Wiltermuth 2011). The comparison between treatments *GroupPC* and *GroupNoPC* will address this argument explicitly. However, there are also good arguments against a stronger prevalence of lying among groups than among individuals or arguments that do not provide a signed prediction. First, social image concerns could be stronger when deciding in a group (see, e.g., Bénabou 2013; Bénabou and Tirole 2006, 2012; Dufwenberg and Dufwenberg 2016).[11] Second, changes in the perception of the norm in the course of the group interaction can drag the comparison between the individual and the group setting toward more or less dishonest behavior. To be conservative, we formulate the following null hypotheses:

H_0^1: *The shares of dishonest reports in part* 2 *in Individual and GroupPC do not differ.*

H_0^2: *The shares of dishonest reports in part* 2 *in GroupPC and GroupNoPC do not differ.*

3. Results

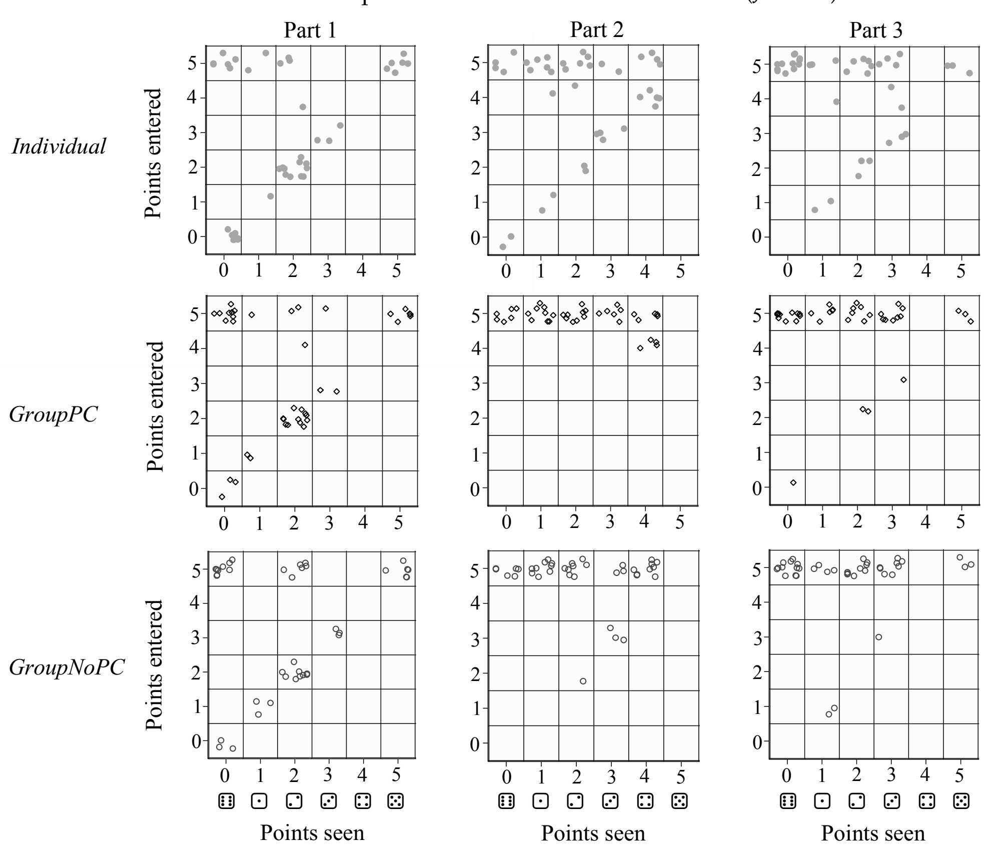
We structure the results section as follows. First, in Section 3.1, we present results on whether groups lie more than individuals, and if so, whether dishonesty becomes more prevalent in groups due to payoff commonality and how group communication affects beliefs about norm compliance (out-of-sample). Then, we analyze the effect of group composition in terms of part 1 liars on group decisions (in Section 3.2). In Section 3.3, we provide a content analysis of the group chat. Section 3.4 provides additional robustness tests for our main results, controlling for personal characteristics of group members.

3.1. Dishonesty Shift and Payoff Commonality

Figure 2 illustrates participants' reporting behavior in parts 1, 2, and 3 by showing the numbers subjects reported (y axis) conditional on the number they have seen (x axis). Note that the figure organizes the axis based on points, which means that the "⚅" is shown at the origin. It is clear that we observe either honest reporting (dots along the 45° line) or dishonest reporting by stating the number that yields the highest returns—i.e., "⚄."[12] The fraction of subjects misreporting in part 1 ranges from 31% to 41% and—as part 1 is identical for all treatments—does not differ significantly between treatments (Fisher's exact tests; *GroupPC* versus *Individual*, $p = 0.810$; *GroupNoPC* versus *Individual*, $p = 0.479$; and *GroupPC* versus *GroupNoPC*, $p = 0.816$). In part 2, we observe significantly more lying in *GroupPC* than in *Individual.* The fraction of dishonest reports amounts to 89.7% in *GroupPC*, whereas in *Individual*, 61.5% of participants misreport their number (Fisher's exact test, $p = 0.007$). The results are qualitatively similar and remain statistically significant if we exclude participants who saw a "⚄" in part 1 (who had no monetary incentive to lie).[13] We thus reject H_0^1.

In *GroupNoPC*, the fraction of dishonest reports amounts to 86.3%, which is also significantly larger than the 61.5% in *Individual* (Fisher's exact test, $p = 0.007$) but does not differ significantly from the median of misreports in *GroupPC* (Fisher's exact test, $p = 1.00$). Hence, we cannot reject H_0^2. We summarize our findings in Result 1.

Figure 2. Die Numbers Seen and Die Numbers Reported Across Treatments and Parts (Jittered)



Notes. In *Individual*, each dot represents one participant's reported number. In the group treatments, each dot represents the median number reported in a group. For readability, the figure shows points (i.e., a die role of "6" is shown as zero points).

Table 1. Coordination Among Group Members in Part 2

	Coordination	Number of dishonest reports	Simulated groups	*Group PC*	*Group NoPC*
Lying decision of individuals in part 2	Yes (all honest)	0	2	4	3
	No	1	11	0	1
	No	2	17	0	5
	No (all dishonest)	3	3	0	0
	Yes (all dishonest)	3	6	35	30
		Σ	39	39	39
Coordination rate (%)			20.5	100	84.6

Notes. Simulated groups are based on part 2 decisions in *Individual*. Note that for three simulated groups, all group members lie, but they still do not coordinate (due to *partial* lying by one member).

Result 1. *Groups lie significantly more than individuals, irrespective of payoff commonalities.*

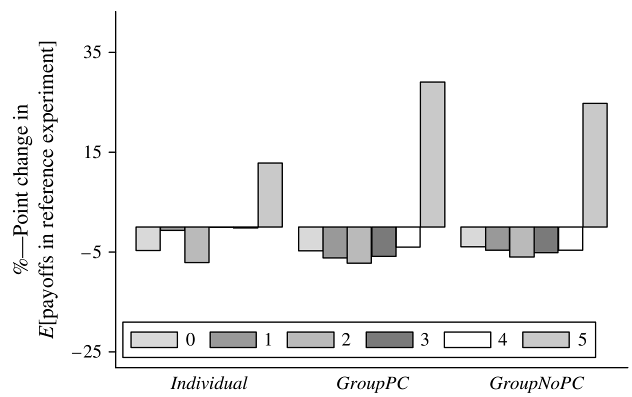
Both payoff commonality and mere communication in *GroupNoPC* strongly foster coordination. In *GroupPC*, all 39 groups coordinate—i.e., all members enter the same number (see Table 1). In *GroupNoPC*, group members do not have to enter the same number to receive a positive payoff and, as compared to *GroupPC*, slightly less members do so (Fisher's exact test, $p = 0.025$). Still, 33 of 39 groups coordinate after the group chat. In five of the six remaining groups, two of three members report dishonestly. To compare coordination in the group treatments with a benchmark, we simulate "coordination" behavior in *Individual* based on participants' actual reporting from part 2 in *Individual*.[14]

As can be seen in Table 1, actual coordination rates in both *GroupPC* and *GroupNoPC* are significantly higher than coordination in simulated groups in *Individual* (Fisher's exact tests, $p < 0.001$ for the two comparisons). We summarize our findings in Result 2.

Result 2. *Communication increases coordination.*

Next, we focus on how the group interaction affects participants' beliefs about reporting behavior of subjects in a reference experiment. We elicited these beliefs once before part 2 and once after part 3. Figure 3 shows how participants' beliefs about reporting behavior of subjects in the reference experiment changed from part 1 to part 3 for the three treatments. In *Individual*, the expected share of subjects reporting a "⚄" in the reference experiment increases from part 1 to part 3 by 12.8 percentage points in *Individual*, whereas the expected fractions of subjects reporting "⚅" and "⚀" decrease. In contrast, when participants interact in a group, their beliefs change to a much larger extent. The expected share of subjects reporting a "⚄" in the reference experiment increases by 28.4 (22.8) percentage points in *GroupPC* (*GroupNoPC*). We thus observe a significant change in the skewness of beliefs on the reported numbers in the reference experiments for group treatments (Wilcoxon signed-rank exact test; *GroupPC*, $p = 0.002$; *GroupNoPC*, $p < 0.001$) but not for *Individual* (Wilcoxon signed-rank exact test, $p = 0.735$).[15] Consequently, the difference in the belief changes is larger in the group treatments (Kolmogorov–Smirnov exact test; *GroupPC* versus *Individual*, $p = 0.090$; *GroupNoPC* versus *Individual*, $p = 0.006$). We conclude with Result 3.

Figure 3. Change in Beliefs About Reported Payoffs in a Reference Experiment in Terms of the Share of Participants That Reported a Payoff $\pi \in \{0, 1, 2, 3, 4, 5\}$

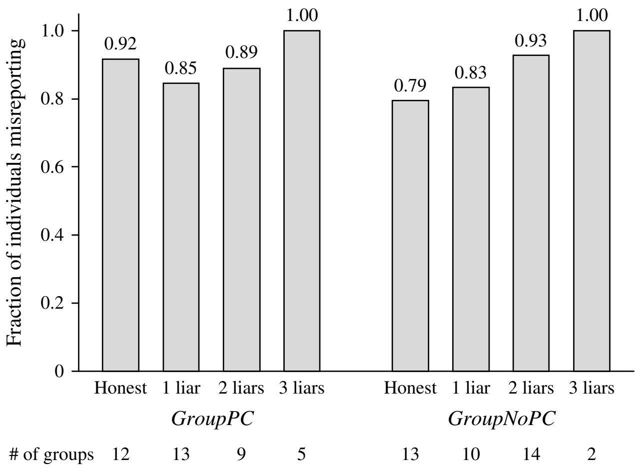


Result 3. *Group communication decreases beliefs about honest behavior of other participants in a reference experiment.*

3.2. The Role of (Dis)Honest Individuals in Groups

We document a strong dishonesty shift in both group treatments casting doubt about the idea that behavior in groups is a simple result of individual preference aggregation due to payoff commonalities. Our experimental design allows us to shed more light on how dishonesty evolves in groups, by studying how the group composition (in terms of participants' individual tendency to report dishonestly in part 1) affects the propensity to lie in part 2. Figure 4 displays the share of group members reporting dishonestly in part 2

Figure 4. Dishonest Reporting Across Treatments Conditional on the Number of Group Members Who Misreported in Part 1



conditional on the number of group members who misreported in part 1. Surprisingly, the group composition does (if at all) play a weak role for dishonest reporting in part 2 of the group treatments (Spearman's rho = 0.135, $p = 0.237$; Fisher's exact test, $p = 0.489$). Even in groups consisting of three previously honest individuals, the vast majority decides to report dishonestly after the group interaction. Only when excluding those participants who have seen a ⚄ in part 1—i.e., those for whom we do not know whether they would have reported dishonestly if they had seen a different number—do we find weak evidence for differences in lying levels across different group compositions (Spearman's rho = 0.253, $p = 0.040$; Fisher's exact test, $p = 0.213$).[16] We summarize this finding in Result 4.

Result 4. *Group composition in terms of lying in part* 1 *affects lying in groups in part* 2 *only weakly. It does not matter for GroupPC, and it matters weakly for GroupNoPC.*

3.3. The Impact of Arguments Used in the Group Chat

Results 1–3 show that communication has a detrimental effect on honest reporting behavior as well as on beliefs about others' honesty. To understand how the group chat affected reporting behavior, four research assistants (to whom the purpose of the study was unknown) independently coded each of the chats using a codebook with a predefined set of variables of interest.[17] Naturally, some coders interpreted the chat protocols differently than others. To obtain reliable values for our variables of interest which, on the one hand, reflect the majority opinion of coders and are, on the other hand, not systematically biased to extreme values, we used the median value for each variable of interest if coders disagreed.[18] We observe on average

Table 2. Number of Groups Using Honest and Dishonest Arguments

Arguments for honesty mentioned	Arguments for dishonesty mentioned: Yes	No	Σ
Yes	13	6	19
No	27	32	59
Σ	40	38	78

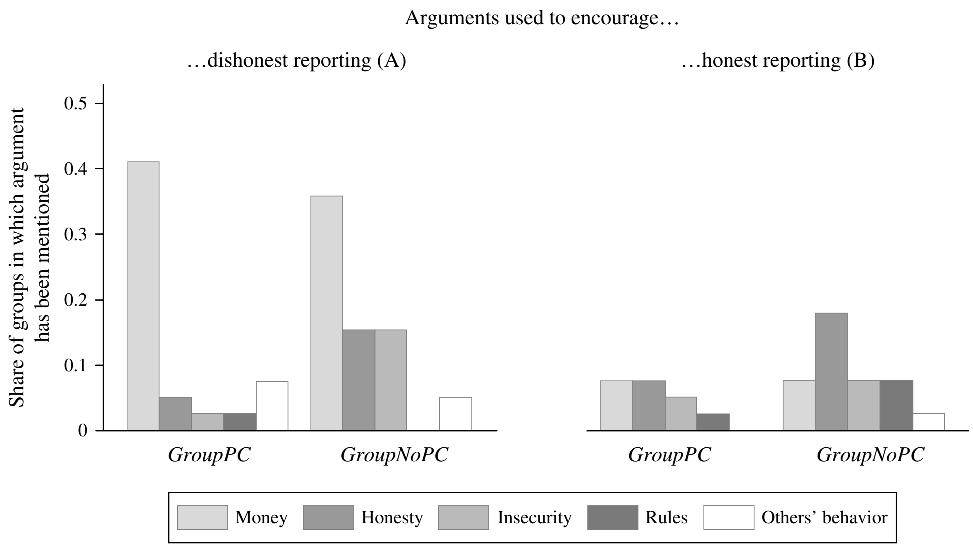
17.45 (std. dev. = 8.76) messages in our 78 groups, and chats last about 162 seconds (std. dev. = 85.7).[19] First, we examine which arguments are used and whether their use is intended to encourage honest or dishonest behavior. Then, we analyze the impact of these arguments on participants' actions.

In 51% of our groups (40 of 78), arguments for dishonesty are explicitly mentioned, whereas only 19 groups make arguments for honesty (see Table 2). Thus, dishonest arguments are made more frequently than honest arguments (χ^2-test, $p = 0.086$). The average share of messages containing arguments for dishonest behavior (i.e., the number of messages using arguments for dishonest reporting in a group divided by the number of messages including arguments for honest and dishonest reporting in the group) amounts to 43.4% and is significantly higher than the share of honest messages (15.6%; within-group comparison, Wilcoxon signed-rank exact test, $p < 0.001$). First suggestions on which number to report are made by both honest and dishonest individuals (classified according to their behavior in part 1) in equal proportion. Comparing whether groups make no argument for (dis)honesty or at least one argument, we do not observe any significant differences between *GroupPC* and *GroupNoPC* concerning dishonest arguments (χ^2-test, $p = 0.365$). More groups without payoff commonality send, however, at least one argument for honesty ($p = 0.065$). We summarize our findings in Result 5.

Result 5. *Arguments for dishonesty occur significantly more frequently than arguments for honesty, irrespective of payoff commonality.*

Let us now turn to what arguments are made in favor of honest and dishonest reporting and how these arguments relate to actual reporting behavior. Figure 5 illustrates the share of groups in which specific arguments were made to encourage dishonest (panel A) and honest reporting (panel B) for *GroupPC* and *GroupNoPC* separately. We refer to *Money* if the argument made related to the monetary consequences of reporting (e.g., "we will earn more if we choose to report a higher number than the number shown"). *Honesty* arguments directly refer to honesty as a norm

Figure 5. Arguments to Encourage (Dis)Honest Reporting



or value (e.g "there is no need to be honest" or "it is important to be honest"). *Insecurity* refers to insecurity concerning the task (e.g., "I am uncertain about the task" or "I thought we should enter the number we want to enter"). *Rules* refer to explicit arguments that include (non)compliance with the rules (e.g., "we should stick to the rules" or "there is no need to stick to the rules"). *Others' behavior* refers to honesty of others outside the group (e.g., people's behavior in general or other participants' behavior).

The majority of arguments refer to *Money, Honesty* and *Insecurity* (see Figure 5). In both *GroupPC* and *GroupNoPC,* participants use *Money* mainly to encourage dishonest reporting, and they do so to a similar extent (Fisher's exact test, $p = 0.233$). Explicitly referring to *Honesty* is the main argument used in favor of honesty, being raised significantly more often in *GroupNoPC* (Fisher's exact test, $p = 0.008$).[20] *Insecurity* is used to encourage both dishonest and honest behavior.

As the use of arguments relating to *Insecurity* does not correlate with wrongly answered control questions in part 2 (Spearman, referring to honesty, rho = −0.05, $p = 0.65$; Spearman, referring to dishonesty, rho = −0.06, $p = 0.59$), participants may use these types of arguments in particular as a justification for their preferred behavior—i.e., as an excuse for dishonest behavior or as support for honest reporting.

We further find that the group composition affects the use of arguments. Groups composed of none or one participant who behaved dishonestly in part 1 use a large variety of arguments (but more dishonest than honest arguments). Groups with two previously dishonest participants use fewer arguments and focus mainly on arguments relating to money. Groups with three previously dishonest people use only arguments relating to money.[21]

Models (1)–(3) in Table 3 regress the probability of a dishonest report on arguments used for encouraging both honest and dishonest behavior. Model (1) shows that, given the high lying rates we observe, arguments that encourage honest reporting explain variation in lying best. We find that arguments in favor of honest reporting that refer to *Money*, *Honesty*, *Insecurity*, and *Rules* reduce the probability of dishonest reporting significantly. In model (2), we introduce the treatment dummy *GroupPC*, which does not significantly influence dishonest reporting (compared to *GroupNoPC*). In model (3), we additionally include whether or not an individual has misreported in part 1. Misreporting in part 1 significantly increases the probability of reporting dishonestly in part 2. Introducing this additional control does not strongly affect the magnitude of the other coefficients. However, the coefficient for *Honesty* (honest use) and *Rules* (honest use) becomes statistically insignificant in model (3).[22] We summarize these findings in Result 6.

Result 6. *Arguments encouraging honesty significantly reduce lying behavior.*

3.4. Individual Characteristics and Lying

In addition to our treatment effects, individual characteristics may play an important role for the decision to report dishonestly. As a robustness test, we report results from a series of probit regressions on dishonest reporting including additional controls for individual characteristics (e.g., Machiavelli scores, big 5, risk attitudes, religiousness, political attitudes, gender, age) in Online Appendix Tables A1–A3. Our treatment effects remain robust, and we find no significant effect of any control variable except for gender.[23] The gender effect that we observe is in line with most previous studies: females tend to be less likely to lie as individuals (see,

Table 3. Lying Behavior and Arguments Used

	(1)	(2)	(3)
		Probit (ME) Misreporting in part 2	
Money (dishonest use)	0.0629	0.0631	0.0572
	(0.0549)	(0.0560)	(0.0532)
Money (honest use)	−0.291***	−0.291***	−0.320***
	(0.131)	(0.131)	(0.134)
Honesty (dishonest use)	0.0437	0.0429	0.0475
	(0.0984)	(0.0949)	(0.0845)
Honesty (honest use)	−0.0773*	−0.0777*	−0.0526
	(0.0382)	(0.0401)	(0.0353)
Insecurity (dishonest use)	0.0187	0.0179	0.00132
	(0.0763)	(0.0750)	(0.0532)
Insecurity (honest use)	−0.375***	−0.376***	−0.343***
	(0.125)	(0.131)	(0.126)
Rules (dishonest use)			
Rules (honest use)	−0.266**	−0.268**	−0.215
	(0.132)	(0.138)	(0.133)
Others' behavior (dishonest use)	−0.162	−0.163	−0.127
	(0.151)	(0.151)	(0.149)
Others' behavior (honest use)			
GroupPC		−0.00245	0.00225
		(0.0579)	(0.0545)
Misreporting in part 1			0.0869**
			(0.0364)
Observations	234	234	234
Cluster	78	78	78
Pseudo-/R-squared	0.422	0.422	0.450

Notes. The dependent variable is a binary indicator for whether the individual has misreported in part 2. Arguments relating to others being honest and arguments referring to *Rules* to encourage dishonest reporting have been dropped due to multicollinearity. Robust standard errors clustered on group level in parentheses.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4. Lying and Gender

	Probit (ME)—Misreporting in part 2			
	All (1)	All (2)	Females (3)	Males (4)
GroupPC		0.205***	0.266***	0.116
		(0.0585)	(0.0781)	(0.0744)
GroupNoPC		0.169***	0.231***	0.0763
		(0.0576)	(0.0770)	(0.0745)
Female	−0.0878**	−0.0824**		
	(0.0381)	(0.0366)		
Observations	273	273	167	106
Clustered SE	117	117	100	76
Pseudo-$R2$	0.017	0.079	0.0829	0.031

Note. Robust standard errors clustered on group level in parentheses.
$^{**}p < 0.05$; $^{***}p < 0.01$.

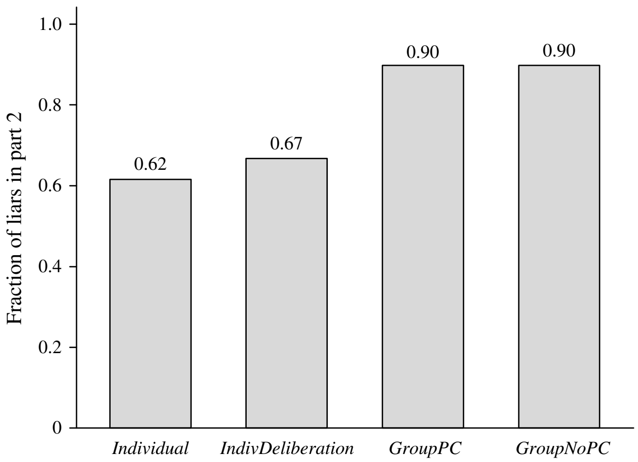
e.g., the survey by Rosenbaum et al. 2014). Regression model (1) in Table 4 shows that females tend to misreport less in part 2. The gender effect persists when controlling for our treatments in model (2). However, the treatment effect on females is larger than on males (see models (3) and (4), showing separate regressions for females and males). The coefficient for males is still positive but not statistically significant at conventional levels ($p = 0.140$ for *GroupPC* and $p = 0.314$ for *GroupNoPC*).

4. Discussion and Alternative Explanations

Our results document that communication in groups can have a detrimental effect on honest reporting behavior. One rationale for this result is that communicating in a group provides possibilities for individuals to deliberate and formulate justifications for dishonest behavior (see also Gino and Ariely 2012, Mazar et al. 2008, Shalvi et al. 2012). Another rationale is that exchanging justifications enables group members to coordinate and establish the validity of a norm regarding honesty and to what extent it is shared among others. To address these two aspects, we conduct an additional control treatment which allows for deliberation but not for communication; *IndividualDeliberation* ($n = 39$). Parts 1 and 3 of *IndividualDeliberation* are identical to the *Individual* treatment described in Section 2. In part 2, *IndividualDeliberation* offers participants the possibility to deliberate before entering the number seen in part 2: after observing the die roll, participants have five minutes to write down their thoughts, and only afterward enter their number. As in the group treatments (where participants could leave the chat), participants in *IndividualDeliberation* were allowed to leave the entry screen before its automated ending. Thus, *IndividualDeliberation* allows participants to deliberate and formulate justifications for dishonest behavior but excludes establishing the validity of arguments and sharing of the norm among group members.

Figure 6 illustrates the fraction of liars in part 2 which is significantly smaller in both individual treatments as compared to the group treatments (Fisher's exact test, $p = 0.026$ for both *IndividualDeliberation* versus *GroupPC* and *IndividualDeliberation* versus *GroupNoPC; p* = 0.007 for both *Individual* versus *GroupPC* and *Individual* versus *GroupNoPC*; $p = 0.814$ for *Individual* versus *IndividualDeliberation*). That is, lying with deliberation is not significantly different from lying in *Individual* but does significantly differ from lying behavior in groups.[24] Also (and again in contrast to group communication), deliberation does not significantly change the skewness of beliefs about other participants' honesty in the reference experiment (Wilcoxon signed-rank exact test, $p = 0.164$). One may expect that individuals in *IndividualDeliberation* deliberate less than individuals in the group treatments. However, the number of individuals in *IndividualDeliberation* coming up with at least one dishonest argument is not statistically different from

Figure 6. Lying in Part 2, by Treatment



the number of groups who do so in *GroupNoPC* (χ^2-test, $p = 0.352$), and it is even larger than the number of groups in *GroupPC* (χ^2-test, $p = 0.068$).[25] Hence, we are confident that deliberation alone does not explain the increase of lying behavior in groups. In combination with Result 1, we conclude that it is the exchange of arguments and moral views within the group that shift group members' expectations and behavior.[26]

Finally, let us briefly discuss other potential explanations for the observed *dishonesty shift*. First, one may suspect that experimenter demand effects could play a role in our setting. For instance, groups might be more susceptible to experimenter demand effects than individuals and thereby lie more (or less). Identifying such demand effects is usually difficult, and there is no existing evidence for a difference in susceptibility between individuals and groups. As we are able to observe communication of our participants, we analyzed group chats with regard to any signs of experimenter demand effects. We find that only 4 of 78 groups (5%) mention that the "honesty–money trade-off" might be part of the research question, and one group considered that the experimenter may be interested in the effect of the group chat on the decision. Also in *IndividualDeliberation*, only 5% (two participants) refer to the research question. Hence, we find little (direct) evidence for an experimenter demand effect that interacts with our group treatments. Such a demand effect is thus unlikely to explain the observed *dishonesty shift.*

Second, we observe increasing lying rates in the *Individual treatment* over the three parts. One may suspect that subjects learn over time that any consequences of dishonest behavior are absent. If so, there is the possibility that communication in groups facilitates this "learning" and thereby increases lying rates faster. However, three empirical facts cast doubt on this interpretation. First, 74% of our groups do not address the potential consequences (positive or negative) of dishonest behavior, and the discussion of potential consequences does not significantly correlate with the group decision to lie (Spearman's rho: positive consequences = −0.069, $p = 0.550$; negative consequences = −0.129, $p = 0.260$). Second, we do not observe a statistically significant correlation between incorrectly answered control question and lying behavior. Third, faster learning in the group treatments in part 2 is unable to explain the significant increase in beliefs about others' dishonesty in a reference experiment that we observe only for participants who communicated in a group. While participants in *Individual* have also learned about the consequences of lying, they do not change their beliefs about others' lying behavior much. Thus, we are confident that group communication indeed changes how people rationalize morally questionable behavior—i.e., it is learning about the norm that mainly drives our results, not about the consequences of lying.

Third, "hiding behind the group" as an individual (accountability argument) is a potential explanation for the dishonesty shift. However, this argument is hard to sustain in an environment with perfect observability by the experimenter, which we implemented on purpose to address this aspect. Further, if hiding behind the group was a main driver, lying in the group treatments should decrease significantly in part 3 (which we do not observe).

5. Conclusion

Our results substantially improve the understanding of (dis)honest behavior in groups. First, complementing recent evidence on the role of collaboration opportunities for dishonesty when communication is absent (see Weisel and Shalvi 2015), we observe that groups lie significantly more than individuals when group members face payoff commonality and have to coordinate on an action. Second, we show that the payoff commonality is not a necessary condition for groups to behave more dishonestly than individuals. If individuals communicate within a group but do not have to coordinate to receive a payoff, their behavior is very similar to the behavior of groups facing payoff commonality. Hence, we provide evidence that communication itself can have a detrimental effect on ethical behavior of small groups. We term this increased inclination to lie *dishonesty shift*. Evidence from our additional control treatment (*IndividualDeliberation*) shows that the *dishonesty shift* is not due to the fact that communicating in a group provides possibilities for individuals to deliberate or simply justify their actions to themselves. Instead, it is the exchange of justifications that enables group members to coordinate on dishonest actions and change their beliefs about moral behavior. The content analysis of our chat protocols backs up this interpretation, as it documents that group members indeed use the chat primarily to formulate and exchange arguments in favor of dishonest behavior. Additionally, we find that communication shifts group members' beliefs about the prevailing honesty norm in a reference experiment, suggesting that group members indeed established a new norm regarding (dis-) honesty.

Our findings also provide important insights for the design of institutions. We document a new argument for the prevalence of dishonesty in groups: groups tend to lie more because they are able to communicate and thereby to rationalize morally wrong behavior in a different way than individuals. In turn, organizations fostering interaction and communication within groups are likely to provide room for group members not only to coordinate their actions, but also to adjust their beliefs and thus the norms they follow. Consequently, members of a group or unit within an organization may hold different beliefs about ethically acceptable behavior than individuals outside of the group. While communication is obviously important and necessary to provide efficient coordination which can be beneficial in many organizational contexts, it may also have detrimental effects on the evolution of norm perception within groups. Therefore, organizational structures that foster communication might require to be paired with strong codes of conduct or exogenous monitoring (and punishment) to avoid the erosion of honesty norms. While reminders appealing to morality or religious beliefs have been shown to be effective in the short run (Mazar et al. 2008), long-run impacts of such interventions are less well understood. More research is also needed to investigate how scrutiny (see, e.g., Baeker and Mechtel 2015, Houser et al. 2016, Van de Ven and Villeval 2015) and efficient reporting mechanisms can be designed (Friesen and Gangadharan 2013). Our results hint at the fact that such designs need to take into account how dishonest behavior evolves in groups and raises many interesting questions for future research—e.g., on the role of hierarchies in organizations, heterogeneity in the benefits from lying, monitoring and punishment, and self-selection of honest and dishonest individuals in leading roles (see, e.g., Fehrler et al. 2016).

Over and above the analysis of potential organizational design features that are able to countervail the dishonesty shift, our results highlight a more general fact: communication in groups and organizations changes group members' beliefs about moral behavior of others and increases coordination of group members' behavior. While we observe a strong shift toward violating the honesty norm in groups, it remains an open question whether group members may coordinate on moral actions, depending on the original strength of the norm.

Acknowledgments

For their valuable comments and suggestions, the authors thank two anonymous reviewers, Johannes Abeler, Alexander Cappelen, Tim Cason, Urs Fischbacher, Anke Gerber, Uri Gneezy, Nicola Lacetera, Gerd Mühlheußer, Daniele Nosenzo, Andreas Roider, Devesh Rustagi, Shaul Shalvi, Ferdinand von Siemens, Matthias Sutter, Bertil Tungodden, Marie-Claire Villeval, Niklas Wallmeier, and Ori Weisel, as well as numerous seminar and conference participants at the Winter Experimental Social Sciences Institute (WESSI) at New York University Abu Dhabi, the Norwegian School of Economics (NHH) Bergen, the University of Chicago, the University of Cologne, the Behavioural and Experimental Economics Group Workshop at the University of Gothenburg, the University of Hamburg, the University of Göttingen, Queensland University of Technology Brisbane, the University of Gothenburg, the London Ph.D. Experimental Workshop (LPEx) at Royal Holloway, Groupe d'Analyse et de Théorie Economique at Lyon, the University of Magdeburg, CAS Munich, the Ludwig-Maximilians-University of Munich, the Technical University of Munich, Evidence-Based Economics Summer Meeting 2016, the University of Nottingham, the University of Melbourne, the Behavioral Spring School 2016 of the University of California at San Diego, the International Association for Research in Economic Psychology/ Society for the Advancement of Behavioral Economics 2016 Meeting, and Tilburg University.

Endnotes

[1] http://www.famous-trials.com/enron/1800-fastowtestimony (last accessed July 27, 2017).

[2] http://www.famous-trials.com/enron/1804-loehrtestimony (last accessed July 27, 2017).

[3] Our variant of the task is computerized and allows the experimenter to observe whether individuals misreport (for a similar idea, see also Gneezy et al. 2016). We explain the variant and the reasons for our design choice in more detail in Section 2.

[4] Note that participants are informed about the experimental procedures. Hence, incomplete lying should occur rarely in our experiment.

[5] As our focus is on how lying behavior changes in groups as compared to individuals (parts 1 and 2), we do not provide a detailed analysis of the data from part 3. Lying rates in part 3 are high (71.8 in *Individual*, 82.1 in *GroupPC*, and 84.6 in *GroupNoPC*) and do not differ significantly across treatment.

[6] This random draw was independent of the random draw that determined payoffs for the lying task.

[7] The translated version of the instructions can be found in Online Appendix C).

[8] As part 3 was identical to part 1, we displayed the control questions in part 3 again but showed the subjects the correct answers. The correlation between the number of times participants failed to answer the control questions correctly and the extent of (dis)honesty is small and fails to be statistically significant (Spearman's rho, individual-level data, part 1: −0.096, $p = 0.112$; part 2: −0.027, $p = 0.663$). Hence, problems in understanding the experimental procedures are unlikely to explain variation in lying behavior between treatments in the experiment.

[9] Two sessions ($n = 39$) of an additional control treatment (*IndividualDeliberation*) that allows us to look into one of the potential explanations were conducted in February 2016 (see also Section 4).

[10] We collapse the group treatment data on the behavior of the group median to ease applicability of nonparametric tests. Using group averages (where applicable) yields very similar results.

[11] Any effects from social image concerns may, however, be less pronounced in our setting, as the group interaction is anonymous, and the experimenter observes individual behavior in all treatments.

[12] The absence of incomplete lying is in line with the fact that there is nothing to disguise in our experiment (in contrast to Fischbacher and Föllmi-Heusi 2013). It also confirms the findings of a similar individual treatment in Gneezy et al. (2016) that uses an "observed game."

[13] As a robustness test, we conducted all analyses reported in this paper also for a subsample that excludes participants who have seen a "⚄" in part 1. All results remain unchanged—i.e., they are qualitatively similar and differences remain significant, except for one result concerning the influence of the group composition on misreporting (which we discuss in Section 3.2). Online Appendix B1 provides more details.

[14] Following the literature on peer effects (Falk and Ichino 2006), we randomly place three different individuals who have seen the same number (i.e., have similar costs of lying) into a group and repeat this procedure until we have generated the same number of groups that have seen the particular number in each of the two group treatments. For this set of groups, we check how many groups coordinate. In total, we simulate 500 sets of groups in this way and take the average over how many groups coordinate in each simulated set.

[15] We calculate the nonparametric skewness $S_{it} = (\mu - \upsilon)/\sigma$ measure for each individual's belief distribution and compare this measure for parts 1 and 3, where μ represents the mean, υ the median, and σ the standard deviation. The tests for the group treatments uses group-level data (collapsed on the median).

[16] For a more detailed analysis, see Online Appendix B1.

[17] The codebook including the complete list of variables can be found in Online Appendix D.

[18] Note that in 83% of cases, at least three coders agreed on a value of a variable. Our results are robust to using average ratings instead of medians.

[19] Chats in *GroupNoPC* are longer than in *GroupPC*, both in terms of messages and duration (Wilcoxon rank-sum exact tests, $p = 0.003$ and $p < 0.001$, respectively). Otherwise we find only minor differences in the structure and content of the chats (as discussed below).

[20] Notice that reporting in part 1 is related to the use of arguments in expected ways: honest behavior in part 1 relates positively to the use of honest arguments in part 2 (Spearman's rho = 0.144, $p = 0.027$), and dishonest behavior in part 1 relates positively to the use of dishonest arguments (Spearman's rho = 0.112, $p = 0.088$).

[21] We thank an anonymous referee for suggesting this analysis.

[22] Adding the number seen in part 2 as an explanatory variable (proxy for the monetary gain from a lie) in models (1)–(3) leaves the results qualitatively unchanged.

[23] Note that other controls do not systematically affect lying behavior in part 1.

[24] Note that levels of lying in part 1 and part 3 in *IndividualDeliberation* (48.7% and 76.9%) do not significantly differ from levels of lying in *Individual* (30.8% and 71.8%; Fisher's exact test; part 1, $p = 0.165$; part 3, $p = 0.796$).

[25] This also holds true for the absolute number of dishonest arguments. Online Appendix B3 provides further results from the content analysis of *IndividualDeliberation*.

[26] This conclusion is further supported by decision times which we analyze in Online Appendix B2.

References

Abeler J, Becker A, Falk A (2014) Representative evidence on lying costs. *J. Public Econom.* 113(May):96–104.

Baeker A, Mechtel M (2015) Peer settings induce cheating on task performance. IAAEU Discussion Paper Series in Economics.

BBC (2016) Russia doping: New Wada report reveals obstructions to testing. (June 16), http://www.bbc.com/sport/olympics/36542577.

Bénabou R (2013) Groupthink: Collective delusions in organizations and markets. *Rev. Econom. Stud.* 80(2):429–462.

Bénabou R, Tirole J (2006) Incentives and prosocial behavior. *Amer. Econom. Rev.* 96(5):1652–1678.

Bénabou R, Tirole J (2012) Laws and norms. IZA Discussion Paper 6290, https://ssrn.com/abstract=1999318.

Cappelen AW, Sørensen EØ, Tungodden B (2013) When do we lie? *J. Econom. Behav. Organ.* 93(September):258–265.

Chytilova J, Korbel V (2014) Individual and group cheating behavior: A field experiment with adolescents. IES Working Paper 2014/06, Institute of Economic Studies, Charles University, Prague.

Conrads J, Irlenbusch B, Rilke RM, Walkowitz G (2013) Lying and team incentives. *J. Econom. Psych.* 34(February):1–7.

Dufwenberg Jr M, Dufwenberg M (2016) Lies in disguise—A theoretical analysis of cheating. Technical report, University of Arizona, Tucson.

Economist, The (2015) The FIFA corruption scandal: Targeting Blatter. (blog by M.V.) (September 25), http://www.economist.com/blogs/gametheory/2015/09/fifa-corruption-scandal.

Erat S, Gneezy U (2012) White lies. *Management Sci.* 58(4):723–733.

Falk A, Ichino A (2006) Clean evidence on peer effects. *J. Labor Econom.* 24(1):38–57.

Falk A, Tirole J (2016) Narratives, imperatives and moral reasoning. Working paper, University of Bonn, Bonn, Germany.

Fehrler S, Fischbacher U, Schneider MT (2016) Who runs? Honesty and self-selection into politics. IZA Discussion Paper 10258, https://ssrn.com/abstract=2849743.

Fischbacher U (2007) z-Tree: Zurich toolbox for ready-made economic experiments. *Experiment. Econom.* 10(2):171–178.

Fischbacher U, Föllmi-Heusi F (2013) Lies in disguise—An experimental study on cheating. *J. Eur. Econom. Assoc.* 11(3):525–547.

Friesen L, Gangadharan L (2013) Designing self-reporting regimes to encourage truth telling: An experimental study. *J. Econom. Behav. Organ.* 94(October):90–102.

Gino F, Ariely D (2012) The dark side of creativity: Original thinkers can be more dishonest. *J. Personality Soc. Psych.* 102(3):445–459.

Gino F, Ayal S, Ariely D (2013) Self-serving altruism? The lure of unethical actions that benefit others. *J. Econom. Behav. Organ.* 93(September):285–292.

Glätzle-Rützler D, Lergetporer P (2015) Lying and age: An experimental study. *J. Econom. Psych.* 46(February):12–25.

Gneezy U (2005) Deception: The role of consequences. *Amer. Econom. Rev.* 95(1):384–394.

Gneezy U, Kajackaite A, Sobel J (2016) Lying aversion and the size of the lie. https://ssrn.com/abstract=2852055.

Gneezy U, Rockenbach B, Serra-Garcia M (2013) Measuring lying aversion. *J. Econom. Behav. Organ.* 93(September):293–300.

Greiner B (2015) Subject pool recruitment procedures: Organizing experiments with ORSEE. *J. Econom. Sci. Assoc.* 1(1):114–125.

Hotten R (2015) Volkswagen: The scandal explained. *BBC News* (December 10), http://www.bbc.com/news/business-34324772.

Houser D, List JA, Piovesan M, Samek A, Winter J (2016) Dishonesty: From parents to children. *Eur. Econom. Rev.* 82:242–254.

Kocher M, Strauß S, Sutter M (2006) Individual or team decision-making—Causes and consequences of self-selection. *Games Econom. Behav.* 56(2):259–270.

Kocher MG, Sutter M (2005) The decision maker matters: Individual versus group behaviour in experimental beauty-contest games. *Econom. J.* 115(500):200–223.

Kröll M, Rustagi D (2016) Got milk? Motivation for honesty and cheating in informal markets: Evidence from India. Research Center SAFE (Sustainable Architecture for Finance in Europe), Goethe University, Frankfurt.

Lundquist T, Ellingsen T, Gribbe E, Johannesson M (2009) The aversion to lying. *J. Econom. Behav. Organ.* 70(1–2):81–92.

Mazar N, Amir O, Ariely D (2008) The dishonesty of honest people: A theory of self-concept maintenance. *J. Marketing Res.* 45(6): 633–644.

Muehlheusser G, Roider A, Wallmeier N (2015) Gender differences in honesty: Groups versus individuals. *Econom. Lett.* 128:25–29.

Pruitt DG, Teger AI (1969) The risky shift in group betting. *J. Experiment. Soc. Psych.* 5(2):115–126.

Rosenbaum SM, Billinger S, Stieglitz N (2014) Let's be honest: A review of experimental evidence of honesty and truth-telling. *J. Econom. Psych.* 45(December):181–196.

Ruiz RR (2015) 4 Cancer charities are accused of fraud. *New York Times* (May 19), http://www.nytimes.com/2015/05/20/business/4-cancer-charities-accused-in-ftc-fraud-case.html.

Shalvi S, Eldar O, Bereby-Meyer Y (2012) Honesty requires time (and lack of justifications). *Psych. Sci.* 23(10):1264–1270.

Sutter M (2009) Deception through telling the truth?! Experimental evidence from individuals and teams. *Econom. J.* 119(534):47–60.

Teger AI, Pruitt DG (1967) Components of group risk taking. *J. Experiment. Soc. Psych.* 3(2):189–205.

Van de Ven J, Villeval MC (2015) Dishonesty under scrutiny. *J. Economic Sci. Assoc.* 1(1):86–99.

Wallace T (2015) How has Volkswagen tricked the US over car emissions? *Telegraph* (September 21), http://www.telegraph.co.uk/finance/newsbysector/industry/11880433/How-has-Volkswagen-tricked-the-US-over-car-emissions.html.

Weisel O, Shalvi S (2015) The collaborative roots of corruption. *Proc. National Acad. Sci. USA* 112(34):10651–10656.

Wiltermuth SS (2011) Cheating more when the spoils are split. *Organ. Behav. Human Decision Processes* 115(2):157–168.