I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15813

Dishonesty as a Collective-Risk Social Dilemma

Shuguang Jiang
Marie Claire Villeval

DECEMBER 2022

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15813

Dishonesty as a Collective-Risk Social Dilemma

Shuguang Jiang
Shandong University

Marie Claire Villeval
Univ Lyon CNRS, GATE and IZA

DECEMBER 2022

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Dishonesty as a Collective-Risk Social Dilemma*

We study cheating as a collective-risk social dilemma in a group setting in which individuals are asked to report their actual outcomes. Misreporting their outcomes increases the individual's earnings but when the sum of claims in the group reaches a certain threshold, a risk of collective sanction affects all the group members, regardless of their individual behavior. Because of the pursuit of selfish interest and a lack of coordination with other group members, the vast majority of individuals eventually earn less than the reservation payoff from honest reporting in the group. Over time, most groups are trapped in a "Tragedy of Dishonesty", despite the presence of moral costs of lying. The risk of collective sanction is triggered less frequently in small groups than in large ones, while priming a collectivist mindset has little effect on lying.

JEL Classification: C92, D01, D91, D62, H41

Keywords: dishonesty, public bad, group size, collectivism, individualism, experiment

Corresponding author:
Marie Claire Villeval
Université de Lyon
GATE UMR 5824
93 Chemin des Mouilles
F-69130 Ecully
France
E-mail: villeval@gate.cnrs.fr

* We are grateful to participants at the Global Online Around-the-Clock Meetings of the Economic Science Association and the China Behavioral and Experimental Economics Forum for very valuable feedback on a previous version of this research. Research assistance from Zhao Lei is gratefully acknowledged. Financial support from the National Social Science Fund of China (21BJL126),Shandong Province Natural Science Foundation (No.ZR2021MG024) and Taishan Scholar Fund (No.tsqn20213022) is gratefully acknowledged. This research has also benefited from the support of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements Avenir (ANR-11-IDEX-007) operated by the French National Research Agency (ANR), and IDEXLYON from Université de Lyon (INDEPTHIDEX/ SBP/2018/03) within the Programme Investissements d'Avenir (ANR-16-IDEX-0005).

1 Introduction

The Enron, Wells Fargo, or WorldCom scandals are examples of organizational failures resulting from corporate fraud by a small fraction of the employees, which had harmful consequences for the entire staff of these companies. There have been several scams in the financial industry resulting in the loss of reputation or closure of some companies, and a general loss of trust in the industry. A 2019 poll conducted by the IPSOS polling institute on a sample of 19,587 adults in 23 countries revealed that only 20% of the respondents considered bankers as trustworthy and 41% considered them as untrustworthy (Skinner and Clemence, 2019).[1] Similar cases of collective sanctions to the whole group due to the misconduct of a few can be found in several areas. In politics, repeat scandals not only damage the incriminated politicians' party in terms of voting intentions but also erode the trust of citizens toward politicians in general (Schwarz and Bless, 1992; Bowler and Karp, 2004; von Sikorski et al., 2020). In the health sector, a handful of individuals not complying with the social distancing guidelines during the COVID-19 pandemic has led to the development of clusters; the multiplication of clusters had dramatic consequences on the well-being and health of entire populations. In sports, the World Anti-Doping Agency commission banned Russia from the 2018 Winter Olympics due to doping and corruption allegations against a fraction of the Russian athletes. These are a few of the examples of how misbehavior among a sufficiently large minority triggers collective sanctions.[2]

These real-world examples suggest that misbehavior could be treated as a social dilemma. Individuals who pursue their selfish interests without considering the negative externalities of their misbehavior may cause the collective failure of members of their communities, including themselves. We propose an approach of cheating as a collective-risk social dilemma that goes beyond the general approach of cheating as a purely individual decision problem or as a group decision without considering collective risks.[3] We focus on three sets of questions. First, when deciding on their moral conduct, do individuals consider the risk that their individual misbehavior may collectively generate or do they simply free-ride on others' honesty? Do they learn over time, in particular from collective losses, to solve the dilemma by moderating their appetite? Second, since it is more diffi-

[1] Other consequences of the securities frauds at WorldCom and Enron include the adoption of restrictive laws, such as the Sarbanes-Oxley Act in 2002 that made it more expensive for small firms to be listed and led to abnormally low small-firm IPO activity in the U.S. since then with further impact on economic growth and employment (Ritter, 2014).

[2] Other anecdotal evidence can be found in the field of transport. In August 2018 a passenger was raped and murdered by her ride-sharing driver from China's largest ride-hailing company, Didi Chuxing, which happened only three months after a similar murder. The two drivers' crimes sparked wide public rage on the safety of the car-hailing service and the platform, and eventually led to the suspension of the carpooling"Hitch" service nationwide. This hurt the interests of more than 10 million drivers and hundreds of millions of passengers who used Didi Hitch.

[3] See reviews by, *e.g.*, Irlenbusch and Villeval (2015); Jacobsen et al. (2018); Gerlach et al. (2019).

cult to downplay one's individual responsibility in smaller groups, does the size of groups affect the way individuals solve the social dilemma? Third, can one influence individual behavior by priming a mindset that should make the group's perspective more salient?

To address these questions, we tested experimentally a cheating game designed as a threshold public bad in which the pursuit of one's selfish individual interest may entail a risk of sanction to the whole group. Players were matched in fixed groups and individually performed a task whose outcome was to be reported to the experimenter. Players could lie and over-report their actual outcome to increase their personal earnings. The task was a variant of the die-in-a-cup paradigm (Fischbacher and Föllmi-Heusi, 2013). In each period, participants were asked to throw a six-sided die in an opaque cup three times and report each outcome, with payoffs increasing in the amount claimed. Since the outcomes are observable only by the player, this task enables lie detection only through the statistical distribution of the reported outcomes. This mimics the frequent situations in which individual misbehavior cannot be detected at a given moment and only excessively high claims at the group level raise suspicion and, possibly, trigger blind sanctions. The game was repeated for 20 periods. Each individual thus rolled the dice 60 times in total.

The novelty is that we introduced a social dilemma via the group setting, turning the die-in-a-cup task from a nonstrategic decision into a strategic game. In each period, if the sum of claims in the group exceeded a given threshold (corresponding to an average claim of 4 per die roll while the expected average truthful claim is 3.5), there was a 60% probability of a collective sanction. If a sanction occurred, all group members lost their payoffs in the given period, regardless of their own reports. This blind sanction mechanism parallels public goods games rule where individuals benefit from the public goods regardless of their individual contribution. These features mimic real-world situations in which detecting misbehavior at the individual level is difficult but the suspicion of a high degree of dishonesty at the group level is sufficient to trigger a collective sanction. In this game, it is optimal for the group that individuals refrain from over-reporting their outcomes to avoid the risk of collective sanction. However, greedy individuals may be tempted to lie to increase their payoff and free ride on others' honesty, which triggers the risk of collective sanction and leads to lower payoffs than those by truthful reporting.

Across four between-subjects treatments, we manipulated two dimensions that may impact individuals' behavior in this social dilemma game. First, we varied the group size: groups consisted of either three or six members. Group size may matter when the risk of collective sanction in a group depends on the aggregate behavior of its members. Indeed, in a small group, each individual is more likely to play a pivotal role; therefore, they may feel a stronger sense of responsibility or guilt for the occurrence of a collective sanction. Moreover, especially when communication is not allowed, it might be easier to coordinate with other members to avoid reaching the critical threshold. These reasons led us to anticipate that individuals would lie less in small groups than in large groups. However, an opposite effect of group size might be observed if individuals are primarily concerned with efficiency. Indeed, if a collective sanction occurs in a large group, the sum of the payoffs that are lost would be higher than that in a small group. Thus, efficiency concerns may lead to more honesty in large groups. It is important to investigate which effect dominates.

The second dimension that we manipulated was the individuals' cultural mindsets, that is, the state of mind and reference frame that could shape the way they think about the decision problem and their within-group interactions. At the beginning of the sessions, we primed participants with either an individualistic or a collectivist mindset, using Goncalo and Staw (2006)'s procedure. The distinction between collectivist and individualistic values has been considered as the main dimension of cultural variations. Individualism increases the extent to which individuals view themselves as independent, whereas collectivism increases the extent to which they perceive themselves as interdependent on others within the group to which they belong (*e.g.*, Hofstede, 1980; Triandis, 2018). Thus, we expected that the priming of the two different mindsets would influence the ability of individuals and groups to cope with the social dilemma.

Previous literature provides ambiguous predictions for our research question. On the one hand, Mazar and Aggarwal (2011) posited that the degree of collectivism that is prevalent in national cultures increases individuals' likelihood to engage in detrimental conduct without violating their own moral standards, through diffusion or displacement of responsibility. In their cross-country study, Gaechter and Schultz (2016) also found that participants in collectivist societies had higher claims, on average, in the die-under-a-cup task than those in individualist societies. Their interpretation was that cultural values influence the prevalence of rule violations in a country, which in turn influences intrinsic honesty. In contrast, studies exploring the differences between collectivism and individualism suggest that individuals in collectivist cultures are, on average, more group-focused (*e.g.*, Oyserman et al., 2002; Kopelman, 2009). They prioritize collective as well as mutual goals (*e.g.*, Dierdorff et al., 2011; Triandis, 2018) and place more value on cooperation (*e.g.*, Oyserman and Lee, 1988; Talhelm et al., 2014), thus making cooperation in social dilemmas easier (Gorodnichenko and Roland, 2011; Marcus and Le, 2013). In our setting, where cheating reflects a social dilemma, we anticipated that priming a collectivist mindset would reduce lying compared to priming an individualist mindset, which would reduce the risk of collective sanction.

Our study provides four main findings. First, we observed a high level of cheating in all treatments: the proportion of individuals whose average claim for their 60 rolls was larger than 4.033 (revealing dishonesty at a 99% confidence level) was 41.67% in small groups under an individualistic mindset and 29.17% in small groups under a collectivist mindset; the respective percentages were 51.39% and 48.61% in large groups. A consequence of such widespread dishonest individual behavior is that groups frequently reached the threshold that triggered the risk of collective sanction. As most liars did not seem to internalize the social consequences of their misbehavior, more than 80% of the participants eventually earned less than the reservation payoff when all group members reported honestly.

The second finding is that the members of small groups claimed, on average, lower numbers than the members of large groups, which resulted in a less frequent risk of collective sanction in small groups. This suggests that a lower diffusion of responsibility, higher guilt for a bad outcome, and/or a higher ability to coordinate prevailed over the effect of efficiency concerns. Our third finding is that priming a collectivist cultural mindset reduced the variance of net payoffs in small groups compared to priming an individualistic mindset. However, its effect on the individuals' claims was imprecisely estimated and it vanished rapidly, showing overall a null impact of priming on behavior.

Finally, the evolution of claims and payoffs over time revealed that experiencing collective sanctions did not teach individuals how to solve the collective-risk social dilemma and could even motivate them to tell more lies to recover their loss. Because they were unable to avoid the risk of collective sanction even with growing experience, individuals and groups became trapped in a "Tragedy of Dishonesty," by analogy with the "Tragedy of the Commons" (Hardin, 1968) in environmental social dilemmas, in which they earned less than if all group members were behaving honestly.

As a point of comparison, we collected the data from an additional treatment in which participants played a threshold public bad game and decided on the amount of their claims without having to roll a die and report the outcome. In this new treatment participants made higher claims than in our original treatment and collective failures occurred significantly more frequently. This comparison reveals the presence of lying aversion and reputational costs of lying when participants were requested to report the outcomes of die rolls. Therefore, what we learn from our study is the following. Analyzing lying as a social dilemma reveals the risk of facing a Tragedy of Dishonesty if the members of a group do not take sufficiently into account the externalities of their behavior on others. This is the case even if the moral costs of lying contribute to reduce the risk of occurrence of collective failures compared to a public bad environment where individuals can make any claim without having to lie. These results suggest several directions for further research, particularly regarding the implications of the social dilemma approach of dishonesty in terms of deterrence mechanisms and possible use of non-monetary levers of action.

The remainder of this article is organized as follows. Section 2 briefly reviews the related literature. Section 3 introduces the experimental design and procedures. Section 4 presents our theoretical predictions. Section 5 develops our experimental results. Finally, section 6 discusses these results and concludes.

2 Literature Review

Our study contributes to three strands in the literature. The first strand relates to cheating in group settings. Studies have investigated the role of team incentives and group environment on cheating (*e.g.*, Sutter, 2009; Conrads et al., 2013; Gino et al., 2013; Chytilova and Korbel, 2014; Muehlheusser et al., 2015; Kocher et al., 2017), and collaborative dishonesty (*e.g.*, Weisel and Shalvi, 2015). They showed that the tendency to cheat is strengthened in groups compared to individual settings. In the absence of payoff commonality, cheating in groups is encouraged by peer effects, the diffusion of responsibility, and a weakening of moral norms, especially when communication is allowed (Kocher et al., 2017). When cheating benefits team members, cheating is more widespread in groups also because of empathy toward group members and because moral concerns are discounted when cheating also benefits others, especially in-groups (Cadsby et al., 2016), which provides a self-serving excuse to lie (Wiltermuth, 2011; Gino et al., 2013).

Extent literature focuses on settings with either payoff independence between players or payoff communality with positive externalities within the group. However, we study cheating when liars' behavior may endogenously generate a risk of cancelling the payoffs of every group member through a collective sanction. Our originality does not lie in the introduction of negative externalities,[4] but in the combination of negative externalities in a lying game with a social dilemma: the potential victims are not passive players but all the members of the group itself. Note that Engel (2015) proposed a theoretical model of dishonesty as a public bad in a study on scientific fraud, but without providing empirical evidence.

Second, by introducing an endogenous sanction mechanism in a stochastic environment, we contribute to the literature on collective-risk social dilemma games (*e.g.*, Milinski et al., 2008; Santos and Pacheco, 2011; Tavoni et al., 2011; Wang et al., 2020). In such games, agents need to coordinate to prevent an undesirable event that would affect them all from occurring, such as anthropogenic climate change. The collective risk is usually introduced in threshold public good games: if the sum of individual contributions does not reach the threshold, then collective damage may hit the group (*e.g.*, Dannenberg et al., 2015). For example, adopting an evolutionary game-theoretical approach, Santos and Pacheco (2011) introduced a risk of collective failure in a repeated contribution game and showed that groups were more successful in coordinating when the risk increased, especially when they were small. We also introduced a collective-risk social dilemma but in a lying game designed as a threshold public bad game in which group members need to coordinate to *avoid* reaching the threshold that triggers a probabilistic sanction. An important difference is that such lying game introduces additional moral costs and

[4]This is also a feature of the sender-receiver games (*e.g.*, Gneezy, 2005), but these games usually involve individual players. Introducing multiple senders in such games increases deception because of a normative shift and a decreased feeling of guilt toward the receiver (Behnk et al., 2017). We differ in that all players in our game have the same role.

considerations in the decision setting.

Third, we contribute to the literature on the diffusion of responsibility, a phenomenon that has been identified in psychology as a source of decreased moral costs in group decisions (Bandura, 2016). Using a threshold public good game in which individuals had to vote to support or reject an immoral action of their group, Rothenhäusler et al. (2018) theoretically demonstrated that pivotality and shared guilt constitute decisive components of the moral costs affecting individuals' choices. Falk and Szech (2013) experimentally demonstrated that a reduced notion of being pivotal in a group, that is, a reduced sense of being decisive for the outcome, results in less moral actions. An individual's willingness to choose a selfish and immoral option decreases with the perceived likelihood of being pivotal (Falk et al., 2020). Consistently with these findings, individuals vote strategically to avoid being pivotal for an unpopular voting result (Bartling et al., 2015) and some of them actively seek an environment to diffuse responsibility and then, make more selfish choices (Brütt et al., 2020). Overall, perceiving less responsibility for the outcome provides individuals with a justification to engage in less moral actions.

We supplement the existing literature on the diffusion of responsibility by testing the effect of group size under different cultural mindsets. Incidentally, this complements previous studies on social dilemmas that found mixed results regarding the effect of group size on cooperation in public good games but in a setting in which moral costs of lying are introduced.[5] Our study also relates to the literature investigating the role of community framing in social dilemma situations (*e.g.*, Liberman et al., 2004; Rege and Telle, 2004; Dufwenberg et al., 2011; Ellingsen et al., 2012). This literature has shown in particular that the community frames, when they have an influence on behavior, play through the coordination of beliefs. We differ from this approach by priming a collective mindset before the game is introduced rather than introducing a social framing in the game itself.

3 Design and Procedures

3.1 Experimental Design

The Game – Our game builds on the die-rolling paradigm developed by Fischbacher and Föllmi-Heusi (2013) to study lying behavior. In each of the 20 periods of the game, participants had to throw a six-sided die in an opaque cup three times and report each outcome on their computer, knowing that each reported point would pay them five tokens. Privacy was ensured, as the outcome of each roll could only be observed by the individual who threw the die. Lying

[5] Some studies found that cooperation declines as the group size increases (*e.g.*, Boyd and Richerson, 1988; Van Huyck et al., 1990; Santos and Pacheco, 2011), whereas others found the opposite or no effect (*e.g.*, Isaac and Walker, 1988; Isaac et al., 1994; Nosenzo et al., 2015; Diederich et al., 2016; Duffy and Xie, 2016; Weimann et al., 2019).

could not be detected at the individual level in a single period but it could be detected statistically, individually and collectively, at the aggregate level. Indeed, participants rolled the die 60 times in total during the game.

The main departure from the original game is that we turned it into a social dilemma game by introducing an endogenous sanction mechanism.[6] At the beginning of the session, participants were randomly assigned to groups of three or six players, depending on the treatment. The composition of the groups was kept fixed throughout the session. Participants were informed that in each period, there could be a risk of collective sanction inflicted on all the members of the group, depending on the sum of the claims in the group. Precisely, the computer program added up all the numbers claimed by the group members in the period and if the sum reached a certain threshold, there was a 60% probability of a collective sanction. This threshold, which was common information, was equal to 12 times the number of group members, corresponding to an average claim of 4 per die roll regardless of the group size, whereas the expected claim was 3.5 under truthful reporting. [7] The consequence of a collective sanction for a group (if it occurred) was to cancel the payoffs of every group member in the period, regardless of their individual claims.

This collective sanction mechanism introduces a social dilemma: over-reporting the die outcomes increases a liar's individual payoff; however, it raises the collective risk of sanction for the whole group. It captures the tension between pursuing one's interest (claiming a higher number than the observed outcome) and serving the collective interest (resisting the temptation to lie to maintain the integrity of the group payoffs). Behaving selfishly in this context raises an additional moral issue because of this social dilemma.

This mechanism has four important features. First, it does not require the identification of dishonesty at the individual level. This is important because in many natural settings proving individual fault is very hard or excessively costly. Second, it is endogenous because it is triggered by the group members' behavior; thus, it involves the individuals' sense of responsibility for the integrity of their group. Third, it involves some uncertainty, as players are unable to communicate to coordinate their actions and avoid hitting the threshold. Even if they reach the threshold, the sanction remains probabilistic. Finally, it is blind because even honest individuals may be punished for the misconduct of others, which could induce feelings of guilt in those tempted to lie.

[6] We also departed from the original die task by asking players to throw the die three times and report each outcome, whereas originally, players were asked to roll the die as many times as they wish but report only the first outcome. Our aim was to collect a sufficient number of observations at the individual level across the 20 periods to detect liars, in the event of a low level of lying, and to detect lies at the group level in each period. We also differed from the original game by using a linear payment scheme, whereas in Fischbacher and Föllmi-Heusi (2013), reporting a "six" paid zero. This was justified by the nature of our sanction mechanism, as explained below.

[7] An average claim of 4 in a group is evidence of lying with a 80% confidence interval.

The game was repeated in 20 periods and participants were paid the sum of their earnings in each period. The vast literature on social dilemmas has identified a typical decay in cooperation over time in the absence of institutions and a rapid move toward the optimum when endogenous sanction mechanisms were introduced. Hence, it was important to test the extent to which our collective sanction mechanism influenced the evolution of honesty over time. At the end of each period, participants received feedback indicating their total claim, whether a collective sanction occurred within the period, and their final payoff within the period. To avoid inducing a feeling of scrutiny, no feedback was provided on the group members' claims.[8]

Since the game involves probabilistic sanctions, we elicited the participants' risk preferences at the beginning of the sessions. We used the Eckel and Grossman (2008) method in which participants have to select one lottery among six.[9] No feedback on the outcome of the selected lottery was provided until the end of the session.

Treatments – A 2×2 factorial design was implemented between subjects. One dimension varied the size of the groups. Each small group (SG) consisted of three members and each large group (LG) had six members. Thus, the threshold triggering the risk of sanction was 36 points in small groups and 72 points in large groups. The second dimension intended to manipulate the participants' mindset by exogenously priming an individualistic (INDI) or a collectivist orientation (COLL). We used the priming method of Goncalo and Staw (2006) in which participants had to complete a pre-experimental survey (see section A.4 in Appendix A). Participants primed with the INDI condition had to write nine statements in total, describing themselves and something unique about themselves, and why it is advantageous to stand out from other people. Participants primed with the COLL condition had to write statements about groups to which they belong, why they think they are like most other people, and why it might be advantageous to blend in with a group. This procedure is commonly used in priming individualistic and collectivist worldviews and is consistent with natural cross-cultural findings (*e.g.*, Oyserman et al., 2002; Oyserman and Lee, 2010; Chatman et al., 2019).

[8]Another reason was to limit peer effects. Based on the literature showing an asymmetric effect of peer information (bad examples have a stronger impact on peers than good ones), we can expect that lying would have been higher if we had provided detailed feedback to the participants on the individual and group claims.

[9]Each lottery had two possible outcomes, each with a 50% probability to be drawn. The payoffs of the six lotteries were: (56, 56), (48, 72), (40, 88), (32, 104), (24, 120), (4, 140). The expected payoff increased from one lottery to the next, as well as its standard deviation. The last two lotteries offered the same expected payoff but the sixth one was riskier.

3.2 Procedures

The experiment was conducted in the computer laboratory of Zhejiang University of Finance and Economics, China, in June 2019, using z-Tree (Fischbacher, 2007). We recruited 240 students from various disciplines by distributing flyers on the campuses of Zhejiang University and Hangzhou Normal University. A total of 10 sessions were conducted, with 24 participants in each session. We ran two sessions for each condition of the Small Group treatment and three sessions for each condition of the Large Group treatment. Table C1 in Appendix C defines the participants' individual characteristics and Table C2 summarizes the sessions and these characteristics. There were no significant differences in these characteristics across treatments (see pairwise treatment comparisons in Table C3 in Appendix).

Upon arrival, participants were randomly assigned to a terminal and were given a set of instructions (see section A.1 in Appendix A). The experimenter read the instructions aloud and answered questions privately. Participants had to answer control questions (see section A.3 in Appendix A) after they had read the instructions. We did not proceed until all participants had answered all questions correctly. Participants performed the risk preference task first, then proceeded to the priming task before playing 20 periods of the die-rolling game. At the end of each session, participants performed the Triad task of association of Talhelm et al. (2014) that aims to measure analytical *vs.* holistic thinking, being respectively more characteristic of individualistic *vs.* collectivist environments. We administered the test of moral identity developed by Aquino and Reed (2009) and measured the participants' willingness to help a group in need in a hypothetical scenario that simulates a social dilemma. We finally recorded a few socio-demographic characteristics (see section A.5 in Appendix A).

The duration of each session was approximately 90 minutes. On average, participants earned RMB 46.09 (Std. Dev. = 9.49, Max = 75, Min = 21.85) (USD 11 in purchasing power parity), including a show-up fee of RMB 5. Earnings were paid privately via Alipay, a third-party mobile/online payment platform, by an assistant who was not aware of the content of the experiment (this information was provided in the instructions).

4 Predictions

Individuals who did not roll a "6" have a financial incentive to claim a higher number than the actual outcome although they are requested to report their actual outcome. Taken in isolation, income maximizers with standard preferences should report 18 after their three rolls. If each group member claims 18, the probability of a collective sanction is 60% in a period and thus, a player's expected payoff is 7.2 points (that is, $18 \times (1 - 60\%)$) for both group sizes. This is lower than the expected payoff from honest reporting, both in small groups (9.11 points) and in large groups (9.74 points).[10] Therefore, from a collective point of view, it would be more efficient not to lie in full because the risk of sanction reduces the expected payoff. However, from an individual perspective, each player has an incentive to deviate from honest reporting if the other group members are reporting honestly. If each player pursues their individual interest without any moral cost of lying, the level of lying in the group triggers the collective sanction mechanism, which lowers payoffs compared with expected earnings under honest reporting. This captures the "Tragedy of Dishonesty" that echoes the "Tragedy of the Commons" in public goods games.

However, the empirical frequency of lying and collective sanctions may deviate from this prediction for several reasons. First, our sanctioning mechanism is probabilistic. If individuals are risk-averse or if they overestimate the other group members' claims, they may moderate their own claims to reduce the risk of a collective sanction. Second, if individuals have social preferences and care about their group members' fate, they may also refrain from claiming the highest outcomes to avoid reducing others' expected payoffs. Third, the previous literature has shown that in individual settings, due to an intrinsic preference for honesty and reputational costs of lying, not all individuals are willing to lie and not all liars lie in full (*e.g.*, Fischbacher and Föllmi-Heusi, 2013; Gneezy et al., 2018; Dufwenberg and Dufwenberg, 2018; Abeler et al., 2019). The moral costs of lying may even be higher in our game if players fear to be responsible for triggering the collective sanction. This may reduce the extent of lying. The third element is specific to our game and introduces a clear motivational difference with standard threshold public bad games in which individuals can freely decide on their claims. This yields our first conjecture:

Conjecture 1. (*Lying and Collective Sanctions)* The pursuit of their selfish interest by individuals triggers collective sanctions. The prevalence of lying and sanctions is reduced by risk attitudes, social preferences, and moral values.

We do not state predictions about the evolution of behavior over time. Indeed, on the one hand, in this social dilemma setting, individuals who initially played the honest strategy may become less willing to cooperate when they learn about the free riding of their group members through the occurrence of a collective sanction. On the contrary, the occurrence of collective sanctions may progressively discourage lying, particularly if sanctions occur several times in a row. We use the data analysis to conclude on the resulting evolution of decisions over time.

[10] In a small group, there are nine die rolls in total in a period. The probability that the sum of the actual outcomes of the nine rolls reaches at least 36 points is 22.04%. In a large group, there are 18 die rolls in total in a period. The probability that the sum of the actual outcomes of the 18 rolls reaches at least 72 points is 12.13%. In expectation, an honest player's claim is 10.5 in both group sizes ($= 3.5 \times 3$). The average expected payoff from honest reporting is thus $10.5 \times (1 - 22.04\% \times 60\%) = 9.11$ points in small groups and $10.5 \times (1 - 12.13\% \times 60\%) = 9.74$ points in large groups.

We expect an effect based on group size for three reasons. First, in small groups, individuals are more likely to deliberate on their pivotal role in triggering the mechanism of collective sanction. Since the probability to reach the threshold of collective sanction with fully honest reporting is 22.04% in small groups and 12.13% in large groups (see footnote 10), an individual's probability to make his or her group reach the threshold is higher in small groups than in large groups for the same total reported number. Second, in small groups with incomplete information on others' claims, liars can share their potential guilt with fewer people than in large groups. Third, coordination without communication to prevent the risk of collective sanction may be slightly less challenging among members of small groups. However,the size of the group may have an opposite impact if individuals have efficiency concerns because in situations of a collective failure, more payoffs are canceled in large groups than in small ones (six players's payoffs vs. three). We assume that the global effect of pivotality, shared guilt, and coordination ability outweighs the efficiency concern. Therefore, we expect individuals to have lower claims in the SG than in the LG treatment. This yields our second conjecture:

Conjecture 2. (*Group Size*) *Individuals claim lower numbers on average in small groups than they do in large groups.*

Our design aimed at priming either an individualistic or a collectivist mindset between subjects. An individualistic mindset is expected to highlight personal interests relative to the group values, whereas a collectivist mindset is expected to favor group perspective taking. Thus, players assigned to the collectivist priming may be more attentive toward their responsibility in the fate of their group and, thus, show a lower propensity to lie, compared with players that received the individualistic priming. This conjecture builds on the previous literature that showed that individuals in collectivist cultures tend to prioritize mutual goals (Oyserman and Lee, 1988) and cooperate more in social dilemmas (Gorodnichenko and Roland, 2011; Marcus and Le, 2013). Moreover, since responsibility is less diffused in small than in large groups, priming a collectivist *vs.* individualistic mindset is expected to have a stronger impact in small than in large groups. This leads to our third conjecture:

Conjecture 3. (*Mindset Priming*) A collectivist mindset reduces average claims compared to an individualistic mindset. The effect is stronger in small groups than in large groups.

5 Results

First, we present a general analysis of claims and the frequency of collective sanctions occurring in groups. Then, we explore the effects of group size and cultural priming on lying behavior. Since individuals' decisions were interdependent in our social dilemma game, the non-parametric analysis was based on the mean decision in the group across the 20 periods, taken as an independent unit of observation (N = 16 in both the SG-INDI and SG-COLL treatments, and N = 12 in both the LG-INDI and LG-COLL treatments). This provides very conservative tests. We complement this analysis with a regression analysis.

5.1 Lying and Frequency of Collective Sanctions

Overall, the average claim for a die roll was 4.11, which is 0.61 points higher than the expected average of 3.5 if reports were truthful. The difference is significant (Wilcoxon signed-rank test, N=56, $p < 0.001$).[11] The first column in Table 1 shows the average claim by treatment. The average claim was significantly higher than 3.5 in each treatment (4.03 in SG-INDI, 3.93 in SG-COLL, 4.23 in LG-INDI, 4.16 in LG-COLL), which suggests that lying was widespread in all conditions.

The average share of high claims (4, 5 and 6) in groups was 66% (63% in the SG-INDI treatment, 63% in SG-COLL, 68% in LG-INDI, and 67% in LG-COLL). Both at the aggregate and the treatment levels, these shares are significantly different from the expected 50% if individuals reported truthfully. Additionally, column (3) in Table 1 indicates that the highest number (*i.e.*, 6) was claimed significantly more frequently than expected from fair die rolls (16.67%) in every treatment; however, it was only marginally significant in SG-COLL).

Table 1: Average and Highest Claims, by Treatment

Treatment	*Average claim* (1)	*p-value* (2)	*Share of '6'* (3)	*p-value* (4)
SG-INDI	4.03 (0.33)	< 0.001***	23.75%	0.037**
SG-COLL	3.93 (0.19)	< 0.001***	19.62%	0.056*
LG-INDI	4.23 (0.31)	< 0.001***	31.18%	0.001***
LG-COLL	4.16 (0.20)	< 0.001***	26.81%	0.001***
TOTAL	4.11 (0.28)	< 0.001***	24.82%	< 0.001***

Notes: SG (LG, resp.) for small (large, resp.) groups. INDI (COLL, resp.) for individualistic (collectivist, resp.) priming. Column (1) represents the average claim over the 20 periods, computed at the group level, with standard deviations in parentheses. Column (2) reports the p-value from Wilcoxon signed-rank tests of the difference between the average claim and 3.5, taking each group over the 20 periods as one independent observation (N=16 in each SG condition and N=12 in each LG condition). Column (3) reports the share of '6' claims, computed at the group level over the 20 periods. Column (4) reports the p-value from Wilcoxon signed-rank tests of the difference between this share and the expected 16.67% in case of truthful reporting. $* \, p < 0.1, ** \, p < 0.05, *** \, p < 0.01$.

[11] All the nonparametric tests reported in the article are two-sided.

The four panels of Figure 1 display the distributions of claims according to treatment. The stars above bars indicate the significance level of the difference between the actual frequency of a particular claim and its predicted frequency in a uniform distribution (16.67%), according to Wilcoxon signed-ranked tests.[12] In all treatments numbers 4 and 6 were claimed significantly more frequently (except number 6 in SG-COLL, for which the difference was only marginally significant) and numbers 1 and 2 were claimed significantly less frequently than expected in a uniform distribution. Furthermore, one-sample Kolmogorov Smirnov tests at the individual level revealed that the distribution of claims differs significantly from the uniform distribution predicted by truthful reporting in each of the four treatments ($p < 0.001$ in each treatment).

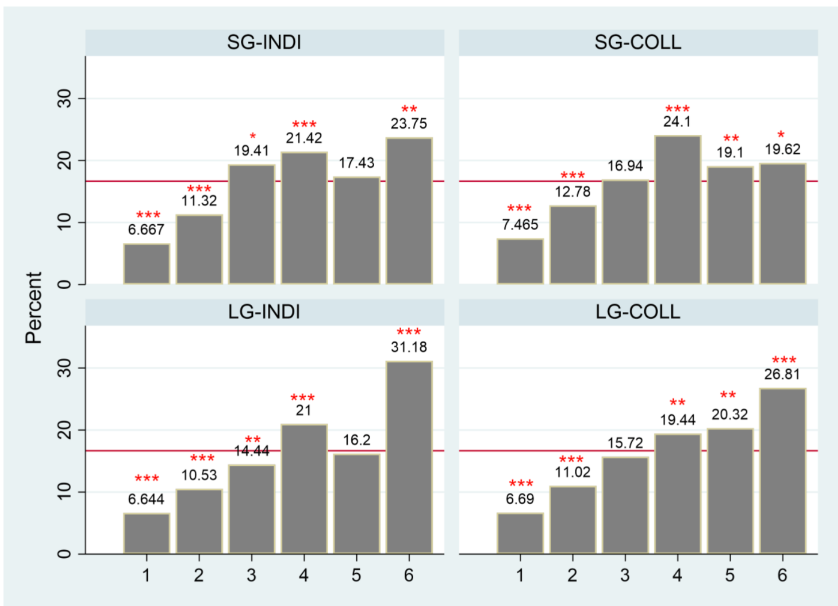


Figure 1: Distribution of the Reports of Die Rolls, by Treatment

Notes. Bars represent the average relative frequency of each number over the 20 periods. Stars above the bars refer to Wilcoxon signed-rank tests at the group level comparing the frequency of a particular report with its expected frequency if individuals reported honestly (16.67%). $* \, p < 0.1, ** \, p < 0.05, *** \, p < 0.01$.

At the individual level, Figure B1 in Appendix B displays the average claim per participant over the 20 periods. Only three players (out of 240) lied to the full extent (*i.e.,* they reported 18 points 20 times). The limited frequency of full lying behavior is consistent with the previous literature in which lying is not directly observable (*e.g.,* Abeler et al., 2019). In small groups, the proportion of individuals whose average claim for their 60 rolls exceeded 4.033 (revealing dishonesty at a 99%

[12]We used Wilcoxon signed-rank tests because the independent observations are at the group level, which requests an analysis at the group level, thus preventing us from using binomial tests.

confidence level) was 41.67% under an individualistic mindset and 29.17% under a collectivist mindset; in large groups, the respective percentages were 51.39% and 48.62%.

We then considered collective sanctions in groups. Table 2 displays the percentage of observations in which the total claims in the group reached the threshold that triggered the risk of a collective sanction (column (1)) and the percentage of observations in which the sanction occurred (column (2)). Table 2 shows that the groups hit the threshold 58% of the time and players lost their payoff 35% of the time because a collective sanction actually occurred. This reveals that most of the time, the groups were unable to coordinate to avoid the risk of sanctions.

Table 2: Relative Frequency of the Risk of Collective Sanction and Actual Sanctions

Treatment	*Frequency of the risk of sanction*	*Frequency of actual sanctions*
SG-INDI	48.13% (0.50)	33.13% (0.47)
SG-COLL	47.19% (0.50)	31.56% (0.46)
LG-INDI	70.83% (0.45)	42.92% (0.50)
LG-COLL	72.92% (0.44)	36.25% (0.48)
TOTAL	58.04% (0.49)	35.45% (0.48)

Notes: These frequencies are computed by reporting the number of periods in which groups reached the threshold (column (1)) or were actually punished (column (2)) to the total number of periods X groups. Standard deviations are in parentheses.

The sanctions triggered by widespread cheating resulted in the participants' average net payoff per period being 37.48 tokens[13] (Std. Dev. = 11.09) in the SG-INDI treatment and 38.45 tokens (Std. Dev. = 6.93) in the SG-COLL treatment. Both were significantly lower than the expected payoff of full honesty in small groups (45.6 tokens, $9.11 points \times 5$) (Wilcoxon signed-rank tests, with one value per group: p = 0.005 in SG-INDI and p<0.001 in SG-COLL). The average net payoff per period was 35.06 tokens (Std. Dev. = 19.92) in LG-INDI and 38.84 tokens (Std. Dev. = 8.44) in LG-COLL. Both are significantly lower than the expected payoff of full honesty in large groups (48.7 tokens, $9.74 points \times 5$) ($p = 0.001$ in LG-INDI and $p = 0.007$ in LG-COLL). Figure 2 displays the average net payoff of each participant per period. 84.2% of the participants (79.17% in small groups and 87.5% in large groups) fall behind the average earnings they would have expected to earn if all group members were reporting honestly.

We examined the evolution of claims to test whether collective sanctions led players to report more honestly over time. Figure 3 displays the average claim per die roll by period and treatment. No treatment exhibits a negative trend over

[13]Recall that each reported point paid 5 tokens, except in the case of a collective sanction where payoffs were void.

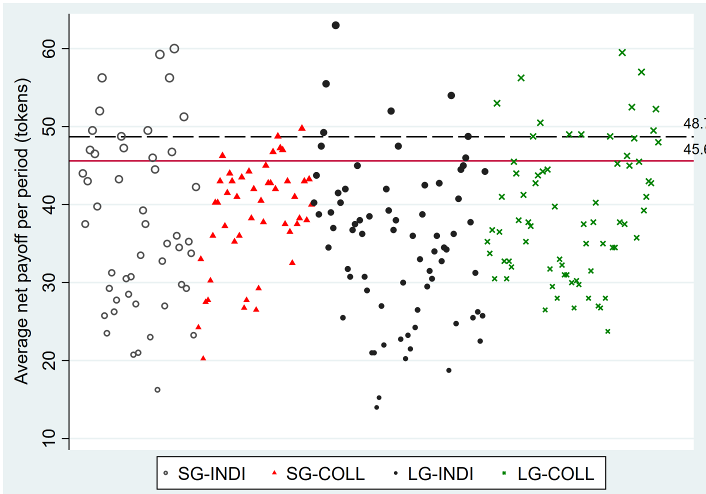


Figure 2: Participants' Average Net Payoff per Period

Notes. The Y-axis represents the average individual payoff per period expressed in tokens. In this figure each dot corresponds to a participant. The horizontal lines correspond to the expected payoff if all group members reported honestly, taking into account the probability that the threshold could be reached if several participants got higher die outcomes by chance.

time. Figure 3 and these statistics show that group members did not learn to solve the social dilemma by cheating less over time in any treatment.

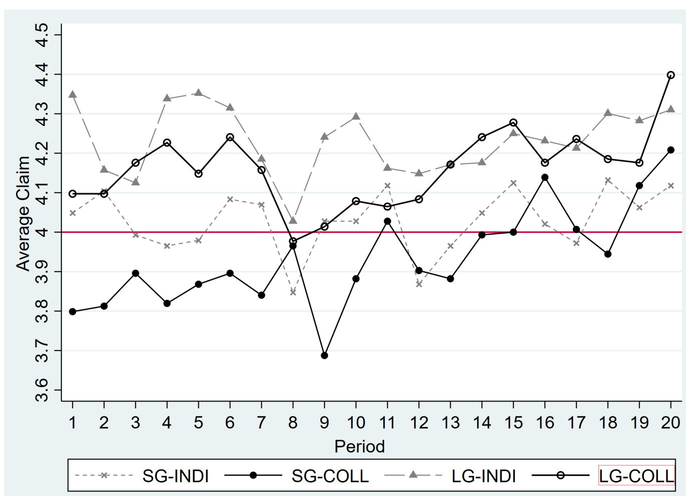


Figure 3: Evolution of the Average Report for a Die Roll over Time

Notes. The Y-axis represents the average number of points reported after a die roll. The X-axis represents the periods. The horizontal line at 4 corresponding to the average threshold triggering the risk of collective sanction in a group.

Result 1 summarizes this descriptive analysis that contributes to support Conjecture 1.

Result 1. ***(Lying and Sanctions)*** *If they rarely lied in full, group members lied sufficiently to frequently trigger the risk of collective sanction. Overall, more than 80% of the players earned less than the expected payoff if all group members were reporting honestly.*

5.2 Effects of Group Size and Mindset Priming

To analyze the impact of group size and mindset priming, we started by examining average claims. As shown in Table 1, claims were higher on average in large groups than in small ones under both priming conditions (this is also visible in Figure B1 in Appendix B which displays the average claim per individual over the 20 periods). Mann Whitney rank-sum tests at the group level indicate that the difference between SG-INDI and LG-INDI ($p = 0.053$) and that between SG-COLL and LG-COLL ($p = 0.006$) are significant (marginally so in the former comparison). However, though the average claim was higher under the individualistic than collectivist priming in both group sizes, the difference between priming conditions is never significant (SG-INDI *vs.* SG-COLL: $p = 0.533$; LG-INDI *vs.* LG-COLL: $p = 0.810$, rank-sum tests). In addition, column (3) in Table 1 shows that the fraction of the highest claim (*i.e.*, 6) was higher in large than in small groups. Mann Whitney rank-sum tests indicate that the differences between SG-INDI and LG-INDI ($p = 0.064$), and between SG-COLL and LG-COLL ($p = 0.005$) are significant (marginally in the former comparison). In contrast, the difference between SG-INDI and SG-COLL ($p = 0.473$) and that between LG-INDI and LG-COLL ($p = 0.434$) are insignificant.

Finally, Mann Whitney rank-sum tests indicate that the percentage of periods in which groups reached the threshold that triggered the sanction mechanism is significantly different between SG-INDI (48.13%) and LG-INDI (70.83%) ($p = 0.045$), and between SG-COLL (47.19) and LG-COLL (72.92%) ($p = 0.003$).[14] In contrast, the difference is insignificant between SG-INDI and SG-COLL ($p = 0.859$), and between LG-INDI and LG-COLL ($p = 0.830$). The higher observed probability of reaching the threshold in large groups is remarkable because in theory, if individuals behaved honestly, this probability would be lower in large groups (12.13%) than in small ones (22.04%). These findings indicate that in large groups players were even less able to coordinate to avoid the risk of collective sanctions. This analysis leads to our second result that supports Conjecture 2:

Result 2. ***(Group Size)*** *The members of small groups claimed lower numbers on average than the members of large groups did, which resulted in a less frequent risk of collective sanction in small groups.*

Although we did not find a significant effect of the nature of priming on the average claim, Table 1 suggests that the level of cheating was higher under the

[14] We conducted this analysis in terms of risk and not in terms of actual sanctions because the occurrence of a sanction depends on the behavior that triggers the risk and on a random draw.

INDI than under the COLL condition. Moreover, SG-COLL is the only treatment in which the average claim was lower than the average threshold value and the number of "6" claims was only marginally significantly higher than expected from truth telling. Figure 3 suggests that this might be driven by a lower level of lying in the first periods in this treatment. To explore differences in claims across treatments over time, Table 3 divides the game into two blocks of 10 periods.

Table 3: Average Claim, by Block of Periods and Treatment

	First 10 periods			*Last 10 periods*		
	SG	LG	*M-W p-value*	SG	LG	*M-W p-value*
INDI	4.015 (0.254)	4.238 (0.244)	0.030**	4.043 (0.450)	4.225 (0.397)	0.142
COLL	3.847 (0.217)	4.121 (0.228)	0.006***	4.022 (0.187)	4.201 (0.217)	0.028**
M-W p-value	0.044**	0.291	-	0.450	0.600	-
sdtest p-value	0.559	0.832	-	0.002***	0.057*	-
N	16	12		16	12	

Notes: The table reports the average claim for a die roll in the first ten periods and in the last ten periods, separately, based on group-level data. Standard deviations are in parentheses. "M-W p-values" are from Mann Whitney rank-sum tests; "sdtest p-values" are from variance-comparison tests. * $p < 0.10$,** $p < 0.05$, *** $p < 0.01$.

Table 3 shows that the average claim in SG-COLL significantly increased from 3.85 points per die roll in the first 10 periods to 4.02 in the last 10 periods ($p < 0.001$, Wilcoxon signed-rank test). The corresponding values in the other treatments are 4.01 and 4.04 in SG-INDI ($p = 0.950$), 4.24 and 4.23 in LG-INDI ($p = 0.791$), and 4.12 and 4.20 in LG-COLL ($p = 0.339$); no significant evolution is found in these treatments. Thus, the difference in average claims between SG-INDI and SG-COLL is significant in the first 10 periods ($p = 0.044$, Mann Whitney rank-sum test), but is insignificant in the last 10 periods ($p = 0.450$). The difference between LG-INDI and LG-COLL is not significant in the first block ($p = 0.291$) or the second one ($p = 0.600$).

Table 3 also shows that while the variance of claims across groups does not differ significantly between priming conditions in the first 10 periods (SG-INDI *vs.* SG-COLL: $p = 0.538$; LG-INDI *vs.* LG-COLL: $p = 0.832$; sd tests), this variance becomes larger in the last 10 periods under the INDI conditions, especially in small groups (SG-INDI *vs.* SG-COLL: $p = 0.002$; LG-INDI *vs.* LG-COLL: $p = 0.057$; sd tests).[15] This specific evolution of claims in SG-COLL is illustrated in a contour plot at the group level in Figure B2 in Appendix B. The contour plot reveals that under the INDI conditions, groups that reported higher (lower, respectively) average numbers in the earlier periods continued to claim higher (lower, respectively) numbers in later periods; in contrast, under the COLL conditions, the contour

[15] As shown in Table 3, the increase of the variance under the INDI conditions in the last 10 periods also makes the group size effect lose significance in this block.

plot is unclear, as group behavior was less stable.

Finally, Figure 2 shows that the variance of the net average payoff per period appears larger in the individualistic priming conditions. Variance-comparison tests conducted at the individual level indicate that the net payoff variance in SG-INDI (Std. Dev. = 11.09) is indeed significantly higher than in SG-COLL (Std. Dev. = 6.93) ($p = 0.002$). The net payoff variance in LG-INDI (Std. Dev. = 9.92) is also larger than in LG-COLL (Std. Dev. = 8.44); however, the difference is not significant ($p = 0.174$). This implies that in small groups, a collectivist mindset priming led to lower income inequality than the individualistic priming.

Overall, these observations suggest the existence of a weak interaction between priming and the group size, namely that the collectivist priming could have had some effect in small groups, as stated in Conjecture 3, but only in the early periods. Over time the social dilemma nature of our game has encouraged free riding in both conditions. This analysis leads to our third result:

Result 3. ***(Mindset Priming)*** *The effects of priming a collectivist cultural mindset on lying were in the expected direction, but they were modest and not robust.*

5.3 Regressions Analysis

In the last section, we report an econometric analysis to support our main results. We estimated linear regressions with robust standard errors clustered at the group level because the players interacted repeatedly within the same group over 20 periods. Table 4 reports marginal effects. The dependent variable is the total claim by each individual for the three die rolls in a period. Models (1) and (2) were estimated on the data from the first 10 periods, models (3) and (4) on the data from the last 10 periods, and models (5) and (6) on all periods. In the models with an uneven number, the independent variables (defined in Table C1 in Appendix C) only include a dummy variable for each treatment (with SG-COLL as the reference category). In the models with an even number, we also included variables that capture risk preferences ("Risk"), moral identity ("Moral Identity"), and willingness to help a group in need ("Help") since Conjecture 1 states that risk attitudes, moral values and social preferences may influence lying behavior. We added socio-demographic variables that control for gender ("Male"), University grade ("Education"), personal monthly income ("Income"), and holistic thinking by opposition to analytical thinking ("Holistic"). The independent variables also include a time trend ("Period") and a dummy variable indicating whether a collective sanction occurred in the previous period ("$\text{Sanction}_{t\text{-}1}$").

Table 4: Determinants of the Individual Total Claim in a Period

Dep. Variable:	*Periods 1-10*		*Periods 11-20*		*All Periods*	
Total Claim	(1)	(2)	(3)	(4)	(5)	(6)
SG-COLL	*Ref.*	*Ref.*	*Ref.*	*Ref.*	*Ref.*	*Ref.*
SG-INDI	0.504**	0.353	0.062	−0.074	0.283	0.104
	(0.245)	(0.282)	(0.357)	(0.323)	(0.277)	(0.273)
LG-INDI	1.174***	0.933***	0.607*	0.334	0.891***	0.603**
	(0.259)	(0.302)	(0.359)	(0.319)	(0.293)	(0.295)
LG-COLL	0.824***	0.697***	0.536**	0.420**	0.680***	0.532***
	(0.249)	(0.235)	(0.228)	(0.186)	(0.214)	(0.181)
Male	-	1.162***	-	1.025***	-	1.091***
		(0.302)		(0.308)		(0.292)
Education	-	−0.020	-	0.013	-	0.001
		(0.086)		(0.085)		(0.081)
Income	-	0.396**	-	0.576***	-	0.490***
		(0.182)		(0.202)		(0.187)
Risk	-	0.289***	-	0.301***	-	0.292***
		(0.079)		(0.085)		(0.076)
Holistic	-	−0.831	-	−0.974*	-	−0.915*
		(0.521)		(0.572)		(0.533)
Moral Identity	-	−0.016	-	−0.016	-	−0.015
		(0.028)		(0.032)		(0.029)
Help	-	−0.091*	-	−0.090	-	−0.089*
		(0.051)		(0.058)		(0.051)
Period	-	−0.018	-	0.053***	-	0.022**
		(0.018)		(0.020)		(0.009)
$\text{Sanction}_{\text{t-1}}$	-	−0.064	-	0.468***	-	0.220**
		(0.128)		(0.120)		(0.095)
Constant	11.540***	10.771***	12.067***	9.797***	11.803***	10.413***
	(0.159)	(0.851)	(0.137)	(0.899)	(0.136)	(0.799)
Nb obs.	2400	2160	2400	2400	4800	4560
Nb groups	56	56	56	56	56	56
R^2	0.022	0.114	0.008	0.119	0.014	0.113
Prob > F	<0.001	<0.001	0.078	<0.001	0.004	<0.001
p values of pairwise tests						
SG-COLL vs. SG-INDI	0.044**	0.216	0.862	0.819	0.311	0.703
LG-COLL vs. LG-INDI	0.216	0.393	0.852	0.797	0.496	0.810
SG-COLL vs. LG-COLL	0.002***	0.004***	0.022**	0.028**	0.002***	0.005***
SG-INDI vs. LG-INDI	0.018**	0.072*	0.250	0.343	0.092*	0.167
SG-COLL vs. LG-INDI	<0.001***	0.003***	0.097*	0.299	0.004***	0.046**

Notes: The table reports the coefficients from OLS regressions with robust standard errors (in parentheses) clustered at the group level. The dependent variable is the sum of the three numbers claimed by a participant in a period. * $p < 0.10$,** $p < 0.05$, *** $p < 0.01$.

The regressions confirm the group size effect identified in the nonparametric analysis in the first 10 periods and when all periods are included for both priming conditions. Supporting Result 2, the sum of claims is significantly higher in large groups than in small ones under the collectivist priming condition (LG-COLL *vs.* SG-COLL: $p = 0.004$ in model (2), $p = 0.028$ in model (4), and $p = 0.005$ in model (6)). Under the individualistic priming condition, LG-INDI also differs significantly from SG-INDI in model (1) ($p = 0.018$, t-test) but the difference fails reaching standard levels of significance in the other regressions. Moreover, behavior in LG-INDI differs from behavior in SG-COLL (LG-INDI *vs.* SG-COLL: $p = 0.003$ in model (2) and $p = 0.046$ in model (6)) (in the second block of periods $p = 0.097$ in model (3) and $p = 0.299$ in model (4)), while this is not the case when we compare SG-INDI with SG-COLL ($p = 0.216$ in model (2), $p = 0.819$ in model (4), and $p = 0.703$ in model (6)). Thus, players cheated more in large groups, but mainly under the collectivist priming condition.

In contrast to group size, most of the time the effect of the mindset priming does not reach standard levels of significance. Model (1) indicates that in small groups, players cheat more in the individualistic mindset than in the collectivist mindset in the first block of periods (SG-COLL *vs.* SG-INDI: $p = 0.044$). However, the significance of the effect disappears once we introduce controls in the regressions (see models (2), (4) and (6)). In large groups, the difference is significant neither in the first block of periods nor when all periods are pooled (t-tests comparing LG-INDI and LG-COLL: $p = 0.393$ in model (2), $p = 0.810$ in model (6)). In the second block of periods, behavior in the LG-COLL treatment differs from the other treatments in model (4), but there is no significant difference between LG-INDI and LG-COLL when controls are not introduced (model (3); $p = 0.797$, t-test). Overall, this regression analysis supports Conjecture 2 but does not support Conjecture 3.

Table 4 shows that the time trend was negative and insignificant in the first 10 periods, but positive and highly significant in the last 10 periods. Similarly, the exposure to a collective sanction in the previous period had a negative but insignificant effect in the first 10 periods and a positive and highly significant effect in the last 10 periods. This could be because some players, at the beginning of the game, may have moderated their claims to avoid the collective sanction, but failed to coordinate at the group level with players who had other motivations; such heterogeneity might explain that the negative coefficients are not significant. Then, in the last 10 periods, receiving an additional collective sanction decreased honesty, probably due to the willingness to recover the loss and because failures progressively revealed information about other group members' dishonesty.

While these characteristics did not differ on average across samples (see Table C3), being a male, having more resources, and a preference for risk had a positive effect on lying in all models. Holistic thinking and a higher willingness to help others had the expected negative sign on the reported claims, but the effects are imprecisely estimated. Moral identity had no significant effect.

To complement the previous regression analysis, we finally estimated the determinants of the number of "6" reported by the players in the period (0, 1, 2, or 3). We used an ordered probit model with standard errors clustered at the group level. The results are reported in Table C4 in Appendix C. They largely confirm the effects (or no effects) identified through the previous regressions. We conclude our analysis with the following result:

Result 4. ***(The Trap of Dishonesty)***: *Despite their experience of collective sanctions, individuals did not learn how to solve the collective-risk social dilemma, and groups were progressively trapped in dishonesty. This was aggravated in larger groups. Priming a collectivist mindset could not prevent this from happening.*

6 Discussion

Our main findings show that although individuals were asked to report their actual outcomes, the temptation to lie to increase one's payoffs led to a high risk of group failure. However, is the risk of such failures as high as if people could directly claim a certain level of earnings in a similar group setting? How does behavior in such setting differ from behavior in a standard public bad game in which players receive no particular request (and so have no reason to lie)? To answer this question and better understand the psychological motivations of the players in our main game, we collected the data from an additional treatment based on a threshold public bad game, keeping the features of our original game constant, except for the request of reporting a random outcome and thus, excluding the lying component. We implemented the small group size and collectivist priming configuration because the SG-COLL condition was used as our reference treatment in the previous analyses and it exhibited the lowest level of lying. Specifically, in the SG-COLL No-Lying treatment participants had to make three claims, by choosing numbers between 1 and 6, in each of the 20 periods, instead of being asked to roll a die and report its outcomes. If the sum of the claims in the group of three players exceeded 36, this triggered a 60% risk of a collective sanction, as in the main treatment.[16]

The SG-COLL No-Lying treatment revealed higher claims, with an average of 4.46 (std. dev. $= 0.33$), which is significantly higher than the average report in the SG-COLL treatment ($p < 0.001$). Figure B3 in Appendix B displays the distribution of claims in this treatment, and the results of Mann-Whitney tests comparing the frequency of each number claimed in this treatment with their counterparts in the SG-COLL treatment. The share of high claims (4, 5, and 6) was 76% and the share of "6" was 33.33%, both significantly higher than in the SG-COLL treatment ($p = 0.001$ and 0.017, respectively). As a result, the sum of claims in the groups triggered the risk of collective sanction in 81.25% of the cases and the collective sanction occurred in 43.13% of the cases, both percentages being significantly higher than in the SG-COLL treatment ($p<0.001$ and $= 0.003$,

[16] This treatment involved 48 new participants in two sessions, forming 16 groups in total. For practical reasons, it was conducted in the experimental laboratory of Shandong University. Table C4 in Appendix C5 indicates that the individual characteristics of these participants did not differ from those of the participants in the original SG-COLL treatment in terms of gender, education level, income, risk attitudes, holistic thinking. However, participants in the new treatment reported a lower moral identity and willingness to help.

respectively). Therefore, the average payoffs were lower than in the SG-COLL treatment, although not significantly so (mean = 36.68 tokens, Std. Dev. = 8.20; p=0.305). The evolution of claims over time showed a very similar increasing pattern to that in the SG-COLL treatment (see B4 in Appendix B).

Whereas risk attitudes or social preferences may play a role in both types of social dilemmas, the moral costs of lying when players were asked to report their outcomes only exist in our lying game. The significant differences observed in the way group members addressed the two social dilemma (the one with direct claims and the other with lying opportunity) reveals the presence of moral costs of lying. These moral costs can explain that individuals claimed less in the main game than in a threshold public bad game with direct claims. As a result, they reduced the risk of collective sanctions, although these moral costs were not sufficiently high to eliminate such risk.[17]

The manipulation of the size of groups in our main experiment also revealed that small groups were better able to avoid triggering the risk of collective sanction than large groups, especially in the first half of the game and under a collective mindset. This probably resulted from a higher sense of pivotality and responsibility of individuals in small groups, although the loss of efficiency in case of a collective sanction was much higher in large groups than in small ones. The differences in the level of claims between the SG-COLL No-Lying treatment and the SG-COLL treatment suggests that there is an interaction effect between the moral costs of lying and the sense of pivotality, since the degree of pivotality of each group member was the same in both conditions.

Finally, priming a collectivist rather than an individualistic cultural mindset had almost no impact: it was restricted to small groups at the beginning of the game, and no longer significant when controlling for individual characteristics. Although a collectivist mindset could help change the lens through which individuals perceived their own behavior and the group's outcome, in fact it did not significantly and durably affect behavior. This lack of effect could be attributed to the fact that group members may feel that they are powerless in such a setting. Of course, we cannot exclude that the priming procedure was perhaps not strong enough to significantly change the players' mindsets. An interesting extension would be to reinforce priming at several points in time to highlight the collective interest of the group and test whether this can discourage misbehavior.

[17]We acknowledge that we can only estimate a net effect of lying aversion. Indeed, lying aversion may have also affected the perception of the risk induced by the other group members' lying behavior, a dimension that we did not elicit. Further investigations would be needed to weigh more precisely the role of risk attitudes, social preferences, lying aversion and the reputational costs of lying.

7 Conclusion

In this study we analyzed dishonesty as a threshold public bad in which individuals' cheating in a group generates an endogenous risk of collective sanction. In real life, individual misbehavior in a community, business organization, or society may lead to a collective loss of reputation, or even the collapse of the whole organization. Such collapse may result in damage even to individuals who followed the moral course of action. This may cause investors and customers to withdraw from such organizations, and generate a high degree of frustration among consumers and citizens. We argue that, at the individual level, this situation can be caused by the temptation of cheating to earn more, even when individuals are aware of the associated collective risk. If most individuals follow their selfish interests when the rule is to report a given outcome but they have the possibility to overreport it, there is a risk of a group collapse. Our experiment mimicked this social dilemma by introducing a probabilistic collective sanction mechanism that was triggered if the sum of individual claims in a die-in-a-cup game reached a certain threshold. Similar in spirit to the Tragedy of the Commons in environmental public goods, a Tragedy of Dishonesty typically occurred: cheating led to the frequent occurrence of collective sanctions that entailed a loss of efficiency in many groups whose members eventually earned less than in situations where there would be truthful reporting. Individuals did not learn to coordinate with their partners to avoid this risk; on the contrary, they tended to lie more over time to recover their losses after a collective sanction instead of behaving more honestly.

By comparing behavior in this game with behavior in a threshold public bad game with a similar sanction mechanism but with no reason to lie, we found that the moral costs of lying led to lower individual claims and a lower associated risk of collective failure in the lying game. This shows that on average, when facing the social dilemma, individuals ask less than what they would like to claim if not having to lie. However, the reference is a situation in which individuals are asked to report their actual outcome, not their desired outcome. In such case, despite the presence of such moral costs of lying, individuals are not honest enough on average to avoid the failure of their groups.

We also found that in the presence of moral costs of lying, small groups induced a lower risk of collective sanction than large groups, probably because of a higher sense of pivotality in small groups. This finding implies that in the presence of such a social dilemma, organizing stakeholders into smaller units might limit dishonest behavior. This seems to be a more promising avenue for interventions than priming a collectivist mindset for which we found no significant effect.

Overall, our study shows that when sanctions are not targeted at individual behavior, the threat of collective failures and the moral costs of lying are not sufficient to discourage misreporting and prevent a normative decline. Future extensions could explore the effect of interventions that have been shown to sustain cooperation in public goods games, such as peer punishment, leadership, and communication.

References

Abeler, J., D. Nosenzo, and C. Raymond (2019): "Preferences for truth-telling," *Econometrica*, 87, 1115–1153.

Aquino, K. and I. I. Reed (2009): "The self-importance of moral identity," *Journal of Personality and Social Psychology*, 83, 1423–.

Bandura, A. (2016): *Moral Disengagement: How People do Harm and Live with Themselves*, Worth Publishers, New York.

Bartling, B., U. Fischbacher, and S. Schudy (2015): "Pivotality and responsibility attribution in sequential voting," *Journal of Public Economics*, 128, 133–139.

Behnk, S., L. Hao, and E. Reuben (2017): "Partners in Crime: Diffusion of Responsibility in Antisocial Behaviors," Working Paper 11031.

Bowler, S. and J. A. Karp (2004): "Politicians, Scandals, and Trust in Government," *Political Behavior*, 26, 271–287.

Boyd, R. and P. J. Richerson (1988): "The Evolution of Reciprocity in Sizable Groups," *Journal of Theoretical Biology*, 132, 337–356.

Brütt, K., A. Schram, and J. Sonnemans (2020): "Endogenous group formation and responsibility diffusion: An experimental study," *Games and Economic Behavior*, 121, 1–31.

Cadsby, C. B., N. Du, and F. Song (2016): "In-group favoritism and moral decision-making," *Journal of Economic Behavior & Organization*, 128, 59–71.

Chatman, J. A., L. L. Greer, E. Sherman, and B. Doerr (2019): "Blurred Lines: How the Collectivism Norm Operates Through Perceived Group Diversity to Boost or Harm Group Performance in Himalayan Mountain Climbing," *Organization Science*, 30, 235–259.

Chytilova, J. and V. Korbel (2014): "Individual and group cheating behavior: a field experiment with adolescents," *IES Working Paper*, 06/2014.

Conrads, J., B. Irlenbusch, R. M. Rilke, and G. Walkowitz (2013): "Lying and team incentives," *Journal of Economic Psychology*, 34, 1–7.

Dannenberg, A., A. Löschel, G. Paolacci, C. Reif, and A. Tavoni (2015): "On the Provision of Public Goods with Probabilistic and Ambiguous Thresholds," *Environmental and Resource Economics*, 61, 365–383.

Diederich, J., T. Goeschl, and I. Waichman (2016): "Group size and the (in)efficiency of pure public good provision," *European Economic Review*, 85, 272–287.

Dierdorff, E. C., S. T. Bell, and J. A. Belohlav (2011): "The power of "we": Effects of psychological collectivism on team performance over time," *Journal of Applied Psychology*, 96, 247–258.

Duffy, J. and H. Xie (2016): "Group size and cooperation among strangers," *Journal of Economic Behavior & Organization*, 126, 55–74.

Dufwenberg, M. and M. A. Dufwenberg (2018): "Lies in disguise – A theoretical analysis of cheating," *Journal of Economic Theory*, 175, 248 – 264.

Dufwenberg, M., S. Gächter, and H. Hennig-Schmidt (2011): "The framing of games and the psychology of play," *Games and Economic Behavior*, 73, 459–478.

Eckel, C. C. and P. J. Grossman (2008): "Forecasting risk attitudes: An experimental study using actual and forecast gamble choices," *Journal of Economic Behavior & Organization*, 68, 1–17.

Ellingsen, T., M. Johannesson, J. Mollerstrom, and S. Munkhammar (2012): "Social framing effects: Preferences or beliefs?" *Games and Economic Behavior*, 76, 117–130.

Engel, C. (2015): "Scientific disintegrity as a public bad," *Perspectives on Psychological Science*, 10, 361–379.

Falk, A., T. Neuber, and N. Szech (2020): "Diffusion of being pivotal and immoral outcomes," *Review of Economic Studies*, 87, 2205—-2229.

Falk, A. and N. Szech (2013): "Morals and markets," *Science*, 340, 707–711.

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

Fischbacher, U. and F. Föllmi-Heusi (2013): "Lies in disguise—an experimental study on cheating," *Journal of the European Economic Association*, 11, 525–547.

Gaechter, S. and J. Schultz (2016): "Intrinsic honesty and the prevalence of rule violations across societies," *Nature*, 496–499.

Gerlach, P., K. Teodorescu, and R. Hertwig (2019): "The truth about lies: A meta-analysis on dishonest behavior," *Psychological Bulletin*, 145, 1–44.

Gino, F., S. Ayal, and D. Ariely (2013): "Self-serving altruism? The lure of unethical actions that benefit others," *Journal of Economic Behavior & organization*, 93, 285–292.

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

Gneezy, U., A. Kajackaite, and J. Sobel (2018): "Lying Aversion and the Size of the Lie," *American Economic Review*, 108, 419–53.

Goncalo, J. A. and B. M. Staw (2006): "Individualism–collectivism and group creativity," *Organizational Behavior and Human Decision Processes*, 100, 96–109.

Gorodnichenko, Y. and G. Roland (2011): "Which dimensions of culture matter for long-run growth?" *American Economic Review*, 101, 492–498.

Hardin, G. (1968): "The tragedy of the commons," *Science*, 162, 1243–1248.

Hofstede, G. (1980): *Culture's consequences*, Beverly Hills: Sage.

Irlenbusch, B. and M. C. Villeval (2015): "Behavioral ethics: how psychology influenced economics and how economics might inform psychology?" *Current Opinion in Psychology*, 6, 87–92.

Isaac, R. M. and J. M. Walker (1988): "Group size effects in public goods provision: The voluntary contributions mechanism," *The Quarterly Journal of Economics*, 103, 179–199.

Isaac, R. M., J. M. Walker, and A. W. Williams (1994): "Group size and the voluntary provision of public goods: experimental evidence utilizing large groups," *Journal of Public Economics*, 54, 1–36.

Jacobsen, C., T. R. Fosgaard, and D. Pascual-Ezama (2018): "Why do we lie? A practical guide to the dishonesty literature," *Journal of Economic Surveys*, 32, 357–387.

Kocher, M. G., S. Schudy, and L. Spantig (2017): "I lie? We lie! Why? Experimental evidence on a dishonesty shift in groups," *Management Science*, 64, 3995–4008.

Kopelman, S. (2009): "The effect of culture and power on cooperation in commons dilemmas: Implications for global resource management," *Organizational Behavior and Human Decision Processes*, 108, 153–163.

Liberman, V., S. M. Samuels, and L. Ross (2004): "The name of the game: Predictive power of reputations versus situational labels in determining prisoner's dilemma game moves," *Personality and Social Psychology Bulletin*, 30, 1175–1185.

Marcus, J. and H. Le (2013): "Interactive effects of levels of individualism–collectivism on cooperation: A meta-analysis," *Journal of Organizational Behavior*, 34, 813–834.

Mazar, N. and P. Aggarwal (2011): "Greasing the palm: Can collectivism promote bribery?" *Psychological Science*, 22, 843–848.

Milinski, M., R. D. Sommerfeld, H. J. Krambeck, F. A. Reed, and J. Marotzke (2008): "The Collective-Risk Social Dilemma and the Prevention of Simulated Dangerous Climate Change," *Proceedings of the National Academy of Sciences*, 105, 2291–2294.

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

Nosenzo, D., S. Quercia, and M. Sefton (2015): "Cooperation in Small Groups: The Effect of Group Size," *Experimental Economics*, 18, 4–14.

Oyserman, D., H. M. Coon, and M. Kemmelmeier (2002): "Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses," *Psychological Bulletin*, 128, 3–72.

Oyserman, D. and S. W. S. Lee (1988): "Collectivism v. individualism: A reconceptualisation of a basic concept in cross-cultural social psychology," in *Cross-cultural studies of personality, attitudes and cognition*, ed. by H. Triandis, London: Palgrave Macmillan, 60–95.

——— (2010): "Priming "culture": Culture as situated cognition," in *Handbook of Cultural Psychology*, ed. by S. Kitayama and D. Cohen, New York: Guildford, 255–279.

Rege, M. and K. Telle (2004): "The impact of social approval and framing on cooperation in public good situations," *Journal of Public Economics*, 88, 1625–1644.

Ritter, J. R. (2014): "Re-energizing the IPO Market," *Journal of Applied Finance*, 24, 37–48.

Rothenhäusler, D., N. Schweizer, and N. Szech (2018): "Guilt in voting and public good games," *European Economic Review*, 101, 664–681.

Santos, F. C. and J. M. Pacheco (2011): "Risk of Collective Failure Provides an Escape from the Tragedy of the Commons," *Proceedings of the National Academy of Sciences of the U.S.A.*, 108, 10421–10425.

Schwarz, N. and H. Bless (1992): "candals and the public's trust in politicians: Assimilation and contrast effects," *Personality and Social Psychology Bulletin*, 18, 574–579.

Skinner, G. and M. Clemence (2019): *Global trust in professions. Who do global citizens trust?*, IPSOS.

SUTTER, M. (2009): "Deception through telling the truth?! Experimental evidence from individuals and teams," *The Economic Journal*, 119, 47–60.

TALHELM, T., X. ZHANG, S. OISHI, C. SHIMIN, D. DUAN, X. LAN, AND S. KITAYAMA (2014): "Large-scale psychological differences within China explained by rice versus wheat agriculture," *Science*, 344, 603–608.

TAVONI, A., A. DANNENBERG, G. KALLIS, AND A. LÖSCHEL (2011): "Inequality, Communication, and the Avoidance of Disastrous Climate Change in a Public Goods Game," *Proceedings of the National Academy of Sciences*, 108, 11825–11829.

TRIANDIS, H. C. (2018): *Individualism and collectivism*, Routledge.

VAN HUYCK, J. B., R. C. BATTALIO, AND R. O. BEIL (1990): "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review*, 80, 234–248.

VON SIKORSKI, C., R. HEISS, AND J. MATTHES (2020): "How Political Scandals Affect the Electorate. Tracing the Eroding and Spillover Effects of Scandals with a Panel Study," *Political Psychology*, 41, 549–568.

WANG, Z., M. JUSUP, H. GUO, L. SHI, S. GEČEK, M. ANAND, AND H. J. SCHELLNHUBER (2020): "Communicating sentiment and outlook reverses inaction against collective risks," *Proceedings of the National Academy of Sciences*, 117, 17650–17655.

WEIMANN, J., J. BROSIG-KOCH, T. HEINRICH, H. HENNIG-SCHMIDT, AND C. KESER (2019): "Public-good provision by large groups - the logic of collective action revisited," *European Economic Review*, 118, 348–363.

WEISEL, O. AND S. SHALVI (2015): "The collaborative roots of corruption," *Proceedings of the National Academy of Sciences*, 112, 10651–10656.

WILTERMUTH, S. S. (2011): "Cheating more when the spoils are split," *Behav. and Human Decision Processes*, 115, 157–168.

A Appendix - Instructions

[*Translated from Chinese*]

Your seat number is:_____. Please take your numbered seat.

A.1 General Instructions

Welcome to this experiment. You have already earned 5 Yuan for showing up on time.

Precautions before starting

1. Please follow the experimenters' instructions during the experiment. Please do not touch the computer until you are instructed to do so. Please keep silent and do not communicate with the other participants during the whole experiment. Please now turn off your cellphone or put it on silent mode and then insert it into the envelope on your desk.

2. **Payment:** You will earn money in this experiment, and your final earnings will be determined by your own decisions, the decisions of other participants, and random draws in the experiment. Note that your final earnings from the experiment will be the sum of payoffs from all parts. All payments in the experiment are expressed in tokens. At the end of the experiment, tokens will be converted to Renminbi at the rate:

   $$100 \text{ tokens} = 5 \text{ yuan Renminbi}.$$

3. **Anonymity:** Your decisions and answers will be anonymous during the whole experiment. During the experiment, you will interact with other participants, but you will not know who your "teammates" are, which means that nobody will know who made any specific decisions. Your experimental payoffs plus the show-up fee will be paid to you in private at the end of the experiment by an assistant who is not aware of the content of this experiment.

If you have finished reading the instructions and do not have any questions, please wait quietly. Otherwise, please raise your hand.

A.2 Task Instructions

SG treatment [*The instructions for the LG treatment are similar, except for the elements into brackets*].

The experiment consists of several parts, which are independent of each other. In each part, you will be asked to make one or more decisions. These parts will be carried out in sequence. Please read carefully the instructions for each part.

Experiment 1

In this experiment, you should select the lottery you would like to play among six different lotteries. The six different lotteries are listed below. Please note:

- You must select one and only one of these lotteries.
- To select a lottery, please click on the corresponding box on the computer screen.

Each lottery has two possible outcomes (Event A or Event B) with a 50/50 chance of occurrence. Your payoff in Experiment 1 will be determined by:

- the lottery you selected and
- the event (Event A or Event B) that occurs.

For example, if you select lottery 4 and Event B occurs, you will earn 104 tokens. If Event A occurs, you will earn 32 tokens. For every lottery, each event has a 50% chance of occurring. At the end of the session, the computer program will randomly determine which event will happen. The computer program will randomly draw a number from 1 to 10 to determine this event. If the drawn number is 1, 2, 3, 4, or 5, Event A will occur. If the drawn number is 6, 7, 8, 9, or 10, Event B will occur. Please click on the "confirm" button on your computer screen when you have made your choice. Once you have confirmed, you cannot change your choice.

On your computer screen, you will have to click on the box corresponding to your preferred lottery, as indicated below:

Lottery	Event	Payoff	Chance
1	A	56	50%
	B	56	50%
2	A	48	50%
	B	72	50%
3	A	40	50%
	B	88	50%
4	A	32	50%
	B	104	50%
5	A	24	50%
	B	120	50%
6	A	4	50%
	B	140	50%

Experiment 2

This experiment consists of 20 periods. In each period, you will face the same decision situation. During the experiment, each participant will be assigned an account to record the increase and decrease of income during the whole process. The initial value of the account is 0 tokens. At the end of the experiment, you will receive a payoff that is the sum of your earnings in each period.

Basic decisions:

In this experiment, you earn money by throwing a die within the cup available on your desk. Specifically, you need to shake the cup to roll the die and then look through the hole on the top of the cup, and report the outcome in the computer. Your payoff depends on the outcome of the die, with 5 tokens for each point. For example, suppose you roll the die in the cup, see 1 point upwards and report this outcome on the computer; then, your account balance will be increased by 5 tokens. Suppose you roll the die in the cup, see 3 points upwards, and report this outcome on the computer; then, your account balance will be increased by 15 tokens. Suppose you roll the die in the cup, see 6 points upwards, and report this outcome on the computer; then, your account balance will be increased by 30 tokens. You have to roll the die three times in each period and report the outcome of each roll in the blank space on the computer screen as follows:

Please note that only you can observe the outcome of your die roll; others, including the experimenter, cannot know the true outcome.

Matching groups and collective failure:

At the beginning of this game, you will be divided into groups of 3 [*6*] participants randomly by the computer program. Thus, you will interact with 2 [*5*] other participants in a group.

Although we cannot know whether the number you report is the actual number on the die, it is possible to infer whether there are untrue reports from the aggregate information of each group. If the total claimed number of your group of 3 [*6*] participants in a period is equal to or higher than 36 [*72*], that is, the average reported number of each participant is equal or higher than 4 (3 participants roll the die 3 times, 3 x 3 x 4 = 36 [*6 participants roll the die 3 times, 6 x 3 x 4 = 72*], the computer program will suspect that there may be untrue reports and it will randomly declare that the outcome of this group in this period is invalid with a 60% chance. This mechanism is called "collective failure" in the experiment.

Specifically, the computer program will automatically calculate the total reported number of each group in each period. When the total number of a group is equal to or higher than 36 [*72*], the computer program will randomly draw a number from 1 to 10. If the drawn number is 1, 2, 3, 4, 5, or 6 (60% chance), a collective failure will occur and the outcome of this group will be void, which implies that the payoff of all participants in this group is 0 in this period. If the drawn number is 7, 8, 9, or 10 (40% chance), all participants in this group still get their payoff according to the number they reported.

At the end of each period, the computer screen will display the sum of your three reported numbers and whether a collective failure occurs in this period, as well as your personal payoff in this period. Then, the program will proceed to the next period decision. The game described above repeats for 20 periods and the period information will be displayed on the top left of your screen during the experiment.

Questionnaire

[*Displayed on the screens*]

During or after the above described Experiment 1 and Experiment 2, you will also have to respond to some questions. All the information collected in the questionnaire will be used for the purpose of scientific research and will be kept confidential. You need to enter your sincere information and ideas after careful consideration, and once entered, they cannot be changed. You have to complete the questionnaire before you can get your payoffs.

Payment

[*Displayed on the screens at the end of the session*]

Since you completed all the experimental and survey tasks, your screen will display your final payoff. Please remain seated until our assistant comes to pay you in private through Alipay. After everybody gets paid, the computer screen will be turned off and then, you can leave; thus, only you know your final payoff.

A.3 Control questions for the SG treatment

[*The questions for the LG treatment are similar, except for the elements into brackets*]

Before we start the experiment, you have to answer several questions about the experiment. You will not be allowed to start the experiment until you answer them correctly. You can ask us for help but you are not allowed to copy the answers of others.

1. Suppose the outcomes of the three die rolls in a period are 3, 4, 5, and you report 3, 4, 5; the total number reported by the other participants in your

group is 23 [*59*]; then, what is the probability of a collective failure in this period for your group? ______ Your payoff in this period is ______ tokens.

2. Suppose the outcomes of the three die rolls in a period are 1, 3, 5, and you report 4, 5, 6; the total number reported by the other participants in your group is 30 [*65*]; then, what is the probability of a collective failure in this period for your group? ______ if a collective failure occurs, your payoff in this period is ______ tokens. if a collective failure does not occur, your payoff in this period is ______ tokens.

3. Suppose the outcomes of the three die rolls in a period are 2, 3, 4, and you report 2, 3, 4; the total number reported by the other participants in your group is 32 [*70*]; then, what is the probability of a collective failure in this period for your group? ______ if a collective failure occurs, your payoff in this period is ______ tokens. if a collective failure does not occur, your payoff in this period is ______ tokens.

A.4 Priming of an individualistic or collectivist mindset

[*The individualistic and collectivist priming questionnaires were taken from Goncalo and Staw (2006). The priming questionnaire was directly displayed on the participants' screens, and not described in the written instructions. The priming took place between the risk elicitation task and the die task.*]

- (*Individualistic priming*)

  A- Write three statements describing yourself.

  B- Write three statements about why you think you are not like most other people.

  C- Write three statements about why you think it might be advantageous to "stand out" from other people.

- (*Collectivist priming*)

  A- Write three statements describing describing the groups to which you belong.

  B- Write three statements about why you think you are you are like most other people.

  C- Write three statements about why you think it might be advantageous to "blend in" with other people.

A.5 Post-experimental questionnaire

[*Displayed on the screens*]

Q1. Triad Task (Talhelm et al., 2014) In the following list, among the three things listed together, please indicate which two of the three are most closely related. Please enter the serial number of the two most relevant things into the box on the right space. If you think A and C are the most relevant, enter AC.

1. A. Seagull B. Sky C. Dog
2. A. Black B. White C. Blue
3. A. Doctor B. Teacher C. Homework
4. A. Apple B. Orange C. Pear
5. A. Train B. Bus C. Tracks
6. A. Shoes B. Boots C. Slippers
7. A. Computer monitor B. Antenna C. Television
8. A. Hospital B. Bank C. Cinema
9. A. Carrot B. Eggplant C. Rabbit
10. A. Cloud B. Wind C. Rain
11. A. Panda B. Banana C. Monkey
12. A. Shirt B. Hat C. Pants
13. A. Kite B. Basketball C. Tennis
14. A. Farmer B. Corn C. Bread
15. A. Shampoo B. Hair C. Beard
16. A. Bridge B. Tunnel C. Highway
17. A. Piano B. Violin C. Guitar
18. A. Child B. Man C. Woman
19. A. Postman B. Policeman C. Uniform
20. A. Letter B. Stamp C. Postcard

Q2. Test of Moral identity (Aquino and Reed, 2009) Listed below are some characteristics that may describe a person: Caring, compassionate, fair, friendly, generous, helpful, hardworking, honest, and kind.

The person with these characteristics could be you or it could be someone else. For a moment, visualize in your mind the kind of person who has these characteristics. Imagine how that person would think, feel, and act. When you have a clear image of what this person would be like, answer the following questions.

[*Participants answered by using a 7-point Likert scale, with 1 for "strongly disagree" and 7 for "strongly agree". The letter "R" in parentheses means that the response was reverse-coded.*]

1. It would make me feel good to be a person who has these characteristics.
2. Being someone who has these characteristics is an important part of who I am.
3. I would be ashamed to be a person who has these characteristics. (R)
4. I strongly desire to have these characteristics.
5. Having these characteristics is not really important to me. (R)

Q3. Willingness to help Suppose you are in a community of 100 members. The community is in a hank and faces the risk of a breakup, which will cause every member to lose 200 Yuan. Now you have 1000 Yuan. What is the maximum number you would like to donate to the community for going through difficulties?

[*Participants had to choose one option among 0, 100, 200, 300, ..., 1000.*]

Q4. Using the scale from 1 to 5 below, please indicate how important you think it is to "be the best" in your studies, work, and life in general?

- 1 Not important at all
- 2
- 3
- 4
- 5 Very important

Q5. Trust Please indicate how much you trust people you meet for the first time in general?

1. Trust completely

2. Trust a little
3. Neither trust, nor distrust
4. Do not trust very much
5. Do not trust at all

Q6. In the experiment, how many participants do you think reported numbers dishonestly? Please enter a number from 0 to 100 to indicate the percentage. _____

Q7. In the experiment, if you expect that other participants in your group would report very high numbers, would you choose to report the true number, a higher number or a lower number?

1. a higher number
2. a lower number
3. the true number

Q8. What is your gender?

1. Female
2. Male

Q9. Are you the only child in your family?

1. No
2. Yes

Q10. Were you born in an urban or rural area?

1. Urban
2. Rural

Q11. What is your current university level?

1. Bachelor - First year
2. Bachelor - Second year
3. Bachelor - Third year
4. Bachelor - Fourth year

5. Master - First year
6. Master - Second year
7. Master - Third year
8. Doctoral studies

Q12. What is your current major? _____

Q13. Are you a member of the Chinese Communist Party?

1. No
2. Yes

Q14. On average, how much money do you receive each month from all channels (including your parents)? _____

A.6 Instructions of the SG-COLL No-lying treatment

[*The other parts of instructions were the same as in the main treatment. Here we only report the instructions of the main task.*].

Experiment 2

This experiment consists of 20 periods. In each period, you will face the same decision situation. During the experiment, each participant will be assigned an account to record the increase and decrease of income during the whole process. The initial value of the account is 0 tokens. At the end of the experiment, you will receive a payoff that is the sum of your earnings in each period.

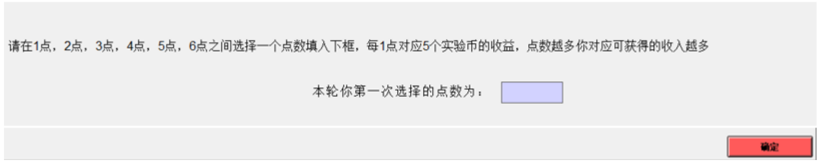
Basic decisions:

In this experiment, you earn money by reporting numbers. Specifically, you need to choose one number among 1 point, 2 points, 3 points, 4 points, 5 points, and 6 points and then report the number on the computer. Your payoff depends on the points you report, with 5 tokens for each point. For example, suppose you report 1 on the computer; then, your account balance will be increased by 5 tokens. Suppose you report 3 on the computer; then, your account balance will be increased by 15 tokens. Suppose you report 6 on the computer; then, your account balance will be increased by 30 tokens. You have to report a number three times in each period, in the blank space on the computer screen as follows:

Please note that only you know the reported number; others cannot know the number.

Matching groups and collective failure:

At the beginning of this game, you will be divided into groups of 3 participants randomly by the computer program. Thus, you will interact with 2 other participants in a group.



Whether you will finally get the payoff corresponding to the numbers you reported also depends on whether the total reported number of your group reached the threshold of "collective failure" and some random factors. Specifically, if the total reported number of your group of 3 participants in a period is equal to or higher than 36, that is, the average reported number of each participant is equal or higher than 4 (3 participants and each reported number for 3 times, 3 x 3 x 4 = 36), the computer program will randomly declare that the outcome of this group in this period is invalid with a 60% chance. This mechanism is called "collective failure" in the experiment.

The computer program will automatically calculate the total reported number of each group in each period. When the total number of a group is equal to or higher than 36, the computer program will randomly draw a number from 1 to 10. If the drawn number is 1, 2, 3, 4, 5, or 6 (60% chance), a collective failure will occur and the outcome of this group will be void, which implies that the payoff of all participants in this group is 0 in this period. If the drawn number is 7, 8, 9, or 10 (40% chance), all participants in this group still get their payoff according to the number they reported.

At the end of each period, the computer screen will display the sum of your three reported numbers and whether a collective failure occurs in this period, as well as your personal payoff in this period. Then, the program will proceed to the next period decision. The game described above repeats for 20 periods and the period information will be displayed on the top left of your screen during the experiment.

B Appendix Figures

Figure B1 displays the average claim over the 20 periods for each participant. Each dot represents one person. The plain horizontal line corresponds to the expected average claim in case of truthful reporting (3.5). The dashed horizontal line corresponds to the average claim that reveals dishonest reporting for 60 rolls at a 99% confidence level (4.033).

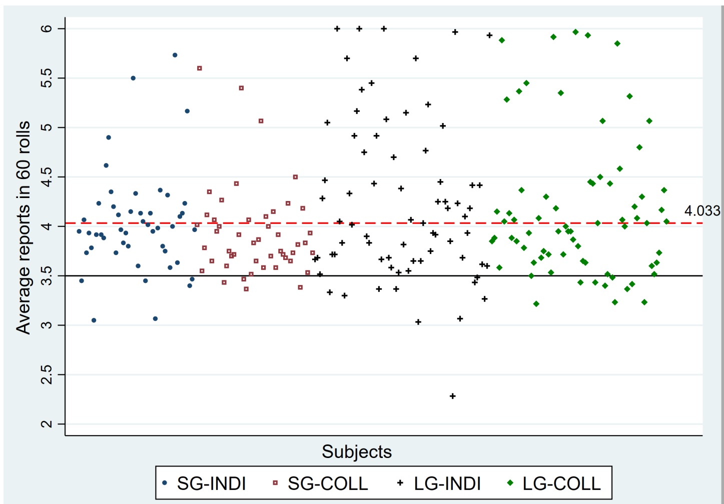


Figure B1: Average Individual Claim

Figure B2 displays contour plots representing the evolution of the average claim in each independent group over the 20 periods of the game. It includes one panel per treatment. Each color represents an interval of average claim. This representation allows us to follow the same group over time and thus, determine whether group behavior was stable or not. The two left panels indicate a relatively stable polarization over time between groups that reported more honestly and groups that lied to a larger extent (they tend to keep the same colors over time). In contrast, the SG-COLL panel shows more variations in the colors at the group level, suggesting that strategies were less stable in these groups.

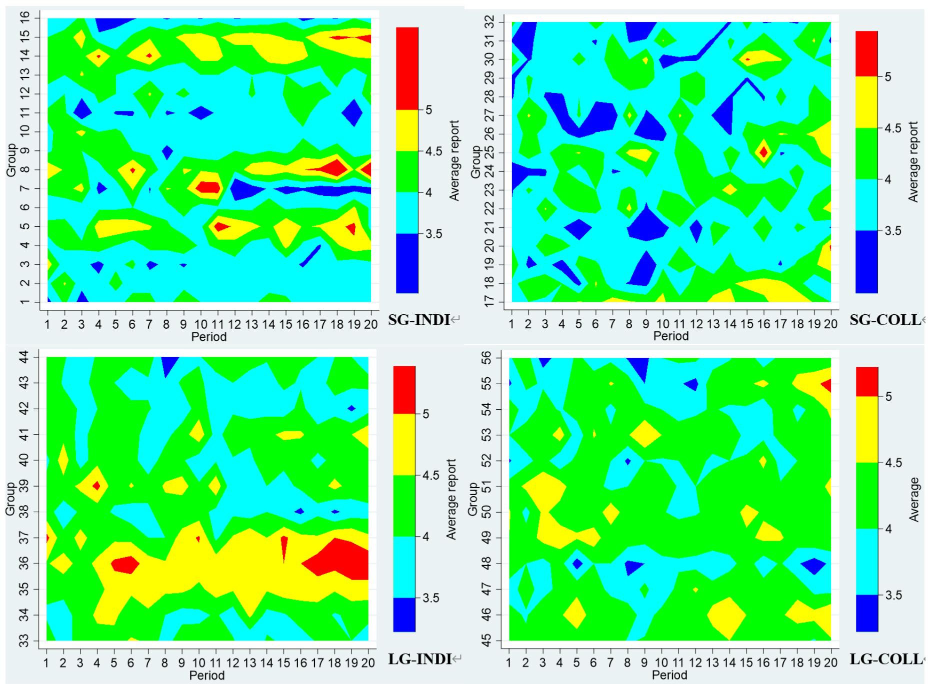


Figure B2: Contour Plot of the Evolution of Each Independent Group's Average Claim Over Time, by Treatment

Notes: The figure represents the average claim in each period per group. Each line on the Y-axis represents one group. The colored scale on the right represents the scope of the average claim (for example, the maximum average group claim was 5.39 in LG-INDI and 6 in SG-INDI, and thus the red bar is lower in LG-INDI than SG-INDI).

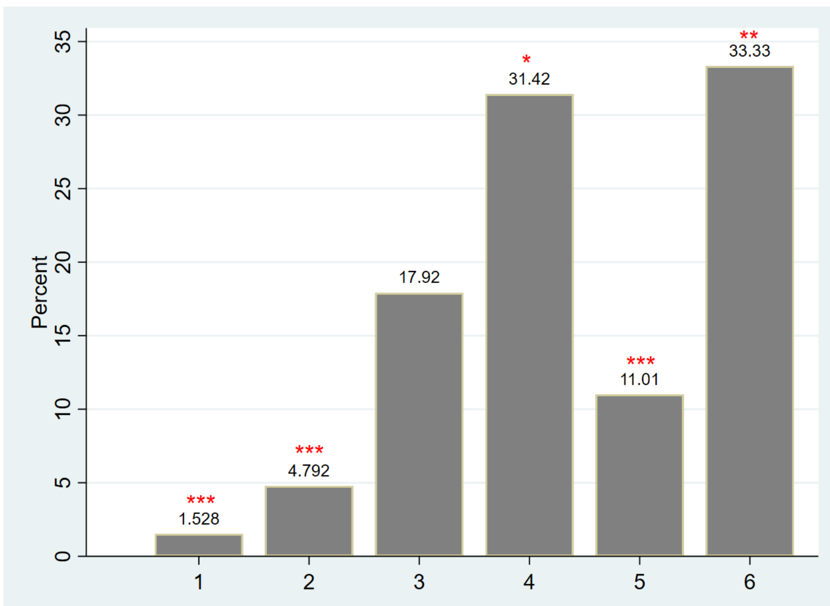


Figure B3: Distribution of Claims in the SG-COLL No-Lying Treatment

Notes. Bars represent the average relative frequency of each number claimed over the 20 periods in the SG-COLL No-Lying treatment. Stars above the bars refer to Mann-Whitney tests at the group level comparing the frequency of a particular claim in the SG-COLL No-Lying treatment with its counterpart in the SG-COLL treatment. $*\ p < 0.10$, $**\ p < 0.05$, $*\!*\!*\ p < 0.01$.

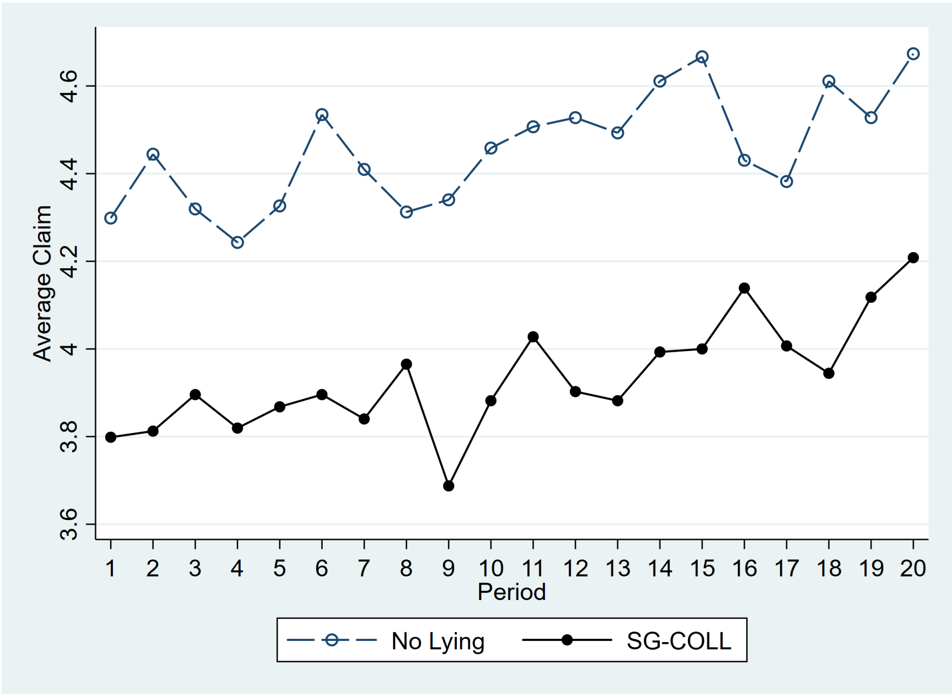


Figure B4: Evolution of the Average Claim Over Time in the SG-COLL No-Lying and SG-COLL Treatments

Notes. The Y-axis represents the average number of points claimed. The X-axis represents the periods.

C Appendix Tables

Table C1: Definition of Variables and Descriptive Statistics

Variables	*Definition*	*Mean*	*Std. Dev.*
Male	Male=1, female=0	0.31	0.46
Education	8 values; Min = 1; Max = 8	2.55	1.56
Income	Monthly income in RMB Value divided by 1000 in regressions	2079	805
Holistic Thinking	Responses in the Triad test (Q1) Min = 0; Max = 1	0.54	0.28
Moral Identity	Sum of responses in the Moral Identity survey (Q2), Min= 7; Max = 35	28.73	4.63
Help	Choice in the Willingness to Help scenario (Q3) 0-10 scale; order ascending with helpfulness	3.61	2.46
Risk	Lottery choice in Part 1. 1-6 scale; ascending order with increasing risk tolerance (5: risk-neutral 6: risk-seeker)	3.91	1.28
$Sanction_{t-1}$	Dummy for whether a collective sanction occurred in the previous period (yes=1, no=0)	0.36	0.48

Table C2: Summary of Sessions and Participants' Individual Characteristics

Treatment	(1) SG-INDI		(2) SG-COLL		(3) LG-INDI		(4) LG-COLL		(5) TOTAL
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	
Male	0.27	0.45	0.27	0.45	0.31	0.46	0.36	0.48	0.31
Education	2.40	1.57	2.56	1.46	2.65	1.55	2.56	1.68	2.55
Income	2082	940	1984	693	2199	911	2021	651	2079
Risk	3.85	1.30	3.85	1.09	4.08	1.25	3.82	1.41	3.91
Holistic thinking	0.47	0.30	0.54	0.28	0.56	0.26	0.56	0.26	0.54
Moral identity	28.31	4.33	29.06	4.97	29.06	4.51	28.44	4.77	28.73
Help	3.56	2.76	3.81	2.27	3.53	2.47	3.58	2.42	3.61
N Sessions	2		2		3		3		10
N Participants	48		48		72		72		240
N Groups	16		16		12		12		56

Notes: The variables are defined in Table C1. SG-INDI for small groups with an individualistic priming; SG-COLL for small groups with a collectivist priming; LG-INDI for large groups with an individualistic priming; LG-COLL for large groups with a collectivist priming. SD for standard deviation.

Table C3: Pairwise Treatment Comparisons of Participants' Individual Characteristics

	p-values					
	(1-2)	(1-3)	(1-4)	(2-3)	(2-4)	(3-4)
Male	1.000	0.682	0.301	0.682	0.301	0.480
Education	0.401	0.267	0.676	0.865	0.657	0.456
Income	0.614	0.756	0.678	0.367	0.881	0.362
Risk	0.836	0.388	0.883	0.260	0.973	0.268
Holistic Thinking	0.227	0.136	0.148	0.928	0.928	0.967
Moral identity	0.241	0.322	0.687	0.755	0.320	0.523
Help	0.374	0.791	0.650	0.358	0.532	0.781

Notes: The table reports the p-values from pairwise treatment comparisons. P-values are from chi-square tests for binary variables and rank-sum tests for interval variables. The numbers in parentheses correspond to the number assigned to the treatment in Table C2: (1) for SG-INDI, (2) for SG-COLL, (3) for LG-INDI and (4) for LG-COLL. The variables are defined in Table C1.

Table C4: Determinants of the Number of "6" Claimed in a Period

Dep. Variable: *Claim of a "6"*	*Periods 1-10* (1)	(2)	*Periods 11-20* (3)	(4)	*All Periods* (5)	(6)
SG-COLL	*Ref.*	*Ref.*	*Ref.*	*Ref.*	*Ref.*	*Ref.*
SG-INDI	0.222**	0.159	0.064	0.024	0.139	0.079
	(0.107)	(0.120)	(0.139)	(0.135)	(0.111)	(0.113)
LG-INDI	0.444***	0.371***	0.285**	0.201	0.360***	0.274**
	(0.104)	(0.122)	(0.127)	(0.124)	(0.105)	(0.113)
LG-COLL	0.323***	0.257**	0.156*	0.121	0.235***	0.178**
	(0.102)	(0.103)	(0.094)	(0.088)	(0.081)	(0.080)
Male	-	0.413***	-	0.318***	-	0.360***
		(0.107)		(0.119)		(0.105)
Education	-	-0.032	-	0.010	-	-0.009
		(0.028)		(0.029)		(0.026)
Income	-	0.075	-	0.187***	-	0.134**
		(0.056)		(0.068)		(0.059)
Risk	-	0.086***	-	0.130***	-	0.109***
		(0.024)		(0.032)		(0.024)
Holistic	-	-0.264	-	-0.199	-	-0.231
		(0.186)		(0.198)		(0.179)
Moral Identity	-	-0.006	-	-0.010	-	-0.008
		(0.010)		(0.011)		(0.010)
Help	-	-0.010	-	-0.022	-	-0.016
		(0.018)		(0.021)		(0.018)
Period	-	0.009	-	0.023**	-	0.012***
		(0.009)		(0.009)		(0.004)
$Sanction_{t-1}$	-	0.014	-	0.176***	-	0.102**
		(0.052)		(0.060)		(0.045)
Nb obs.	2400	2160	2400	2400	4800	4560
Nb groups	56	56	56	56	56	56
$Pseudo R^2$	0.009	0.033	0.004	0.041	0.006	0.035
$Prob > chi2$	$<$0.001	$<$0.001	0.120	$<$0.001	0.003	$<$0.001

Notes: The table reports the coefficients from Ordered Probit regressions with robust standard errors (in parentheses) clustered at the group level. The dependent variable is the number of "6" claimed by a participant in a period (0, 1, 2, 3). * $p < 0.10$,** $p < 0.05$, *** $p < 0.01$.

Table C5: SG-COLL and SG-COLL No-Lying Treatment Comparisons of Participants' Individual Characteristics

Treatment	(1) SG-COLL No-Lying		(2) SG-COLL		(3) SG-COLL No-Lying *vs.* SG-COLL
	Mean	*SD*	*Mean*	*SD*	*p-values*
Male	0.37	0.49	0.27	0.45	0.275
Education	3.25	2.08	2.56	1.46	0.183
Income	2030	747	1984	693	0.960
Risk	4.06	1.56	3.85	1.09	0.399
Holistic thinking	0.53	0.27	0.54	0.28	0.635
Moral identity	26.83	5.49	29.06	4.97	0.022
Help	2.50	1.97	3.81	2.27	0.003
Average claim	4.46	0.47	3.93	0.19	<0.001
Share of "4,5,6"	75.76%	0.13	62.81 %	0.06	0.001
Share of "6"	33.33%	0.19	19.62%	0.07	0.017
Frequency of the risk of sanction	81.3%	0.39	47.19%	0.50	<0.001
Frequency of actual sanctions	43.1%	0.50	31.56%	0.46	0.010
N Sessions	2		2		
N Participants	48		48		
N Groups	16		16		

Notes: The table reports the individual characteristics of the participants in the additional SG-COLL No-Lying treatment and pairwise comparisons with the original SG-COLL treatment. P-values are from chi-square tests for binary variables and rank-sum tests for interval variables. The variables are defined in Table C1.