The title

|  |
| --- |
| Sebastian Simon1 & Jörg A. J. Gross1,2 |
| 1 Social, Economic, and Organizational Psychology, Leiden University |
| 2 Social, Economic, and Organizational Psychology, Leiden University |
|  |

# Author note

Correspondence concerning this article should be addressed to Sebastian Simon, Postal address. E-mail: [my@email.com](mailto:my@email.com)

Abstract

Abstract goes here.

*Keywords:* Rule-following behavior, partner choice, abidance, corruption

The title

Complex social norm systems are rules and standards that guide human behavior in groups (Cialdini, 2001; Pepitone, 1976) and are unique to human beings (Fehr & Fischbacher, 2003; Fehr & Rockenbach, 2004; Gintis, 2003; Ostrom, 2000; Sethi & Somanathan, 1996; Tomasello & Rakoczy, 2003). Importantly, abiding to social norms is generally considered desirable as doing so is associated with good character (Freud, 1977; Hoffman, 1977) which, in turn, helps building and maintaining functional societies [ref]. For example, policy makers and legislators set guidelines and rules for societies that individuals should adhere to and in response to such structures, rule-following behavior then reflects moral character signaling trustworthiness (Baumard, Osiurak, Lesourd, & Le Gall, 2014; Everett, Pizarro, & Crockett, 2016; Freud, 1977; Hoffman, 1977).

Yet, people often choose to violate norms (Köbis, Prooijen, Righetti, & Van Lange, 2016) and engage in corrupt collaborations (Weisel & Shalvi, 2015) such as bribery (International Monetary Fund, 2016; Rose-Ackerman & Palifka, 2016), trafficking (Nieto, 2012), crime and fraud by definition (Ades & Tella, 1996) among many others. The consequences can be fatal including hindrance of economic growth (Mauro, 1995) as well as undermined legitimacy and capacity of governments (Rothstein, 2011), hence, harming the overall functioning of societies. For example, Mexico has been the center of drug violence, killing throusands of people, and exposing the Mexican society to torture and impunity (Nieto, 2012). Also, the Islamic State (IS) caused numerous organized terrorist attacks across Europe over the recent years (Farwell, 2014). We therefore need to determine under which circumstances people abide or break rules.

Previous research highlighted the importance of interpersonal variables in predicting cooperative behavior through morality goals. As people like to think of themselves as moral beings (Abeler, Becker, & Falk, 2014; Abeler, Nosenzo, & Raymond, 2016; Jordan, Mullen, & Murnighan, 2011; Mazar, Amir, & Ariely, 2008) and care about what others think of them (Gausel & Leach, 2011; Lacetera & Macis, 2010; Utikal & Fischbacher, 2013), moral traits dominate (Cottrell, Neuberg, & Li, 2007; Landy, Piazza, & Goodwin, 2016; Landy & Uhlmann, 2018; Peeters, 1992; Wojciszke, Abele, & Baryla, 2009). For example, besides traits such as being sociable and warm (Goodwin, 2015) people prefer to interact with others who signal their trustworthiness (Freud, 1977; Hoffman, 1977) because doing so signals being a ‘good’ interaction partner.

Humans are social beings and when in groups, people prefer others they can rely on, where cues of morality such as rule-following behavior are a good indicator of this (Abele & Wojciszke, 2014; Baumard et al., 2014; Cottrell et al., 2007; Landy et al., 2016; Landy & Uhlmann, 2018; Peeters, 1992; Wojciszke et al., 2009). In this regard, people typically perform costly behaviors that signal trust [ref]. For example, companies in the free market such as Fairtrade and Tony’s Chocolonely use logos with moral appeals to signal their trustworthiness to both their customers and other companies, all of which aim to facilitate cooperation. But companies cannot function in isolation, rather in a market with competitors, and it is this competition that can make companies and individuals more cooperative. For example, Barclay and Willer (2006) found that when people could benefit from being chosen as interaction partners, people engaged in competetive altruism as they were more generous and actively competing against other group members. Hence, with multiple agents, competition pressures arise.

The effects of partner choice have mainly been investigated in cooperation settings using economic games, where providing people with the freedom to seek out trustworthy partners and abandon free-riders has been shown to develop and safeguard cooperation (Efferson, Roca, Vogt, & Helbing, 2016; Rand, Arbesman, & Christakis, 2011) and hence it is typically more lucrative to cooperate than to defect. For example, in hunter-gatherer networks, those who share their spoils form more profitable relationships than those who do not (Gurven, Allen-Arave, Hill, & Hurtado, 2000) and this is no different for nations collaborating with other nations within international alliances such as the European Union (EU). With an increasing number of agents, people prefer to interact with others who are most able and inclined to benefit others (André & Baumard, 2011; Barclay, 2013; Baumard, André, & Sperber, 2013), where benefits can take on the form of wealth, mutual gain, talents, and resources (Hirschman, 1987; Montesquieu, 1951).

Yet, in other settings, people choose to corrupt as a collective, where corrupt collaboration is defined as attaining personal profits as the result of joint acts of rule violations (Weisel & Shalvi, 2015), and depends on the omral goals people have (Baumard et al., 2014; Everett et al., 2016) and the context people are in (Melnikoff & Bailey, 2018). Example for such corruption is the Volkswagen scandal from 2005 where the employees of the company manipulated software to pass key emission tests in the face of limited time and budget and hence likely personal benefits (Goodman, 2015). However, as if corrupt collaboration itself was not enough, corruption breeds corruption. For example, countries that are plagued by more corruption have been found to have a higher black market premium (Bahmani-Oskooee & Goswami, 2005), and higher inflation rates (Cukierman, Edwards, & Tabellini, 1989), all of which go at the expense of the poor (Gupta, Davoodi, & Alonso-Terme, 1998). As these examples demonstrate, we have to research the determinants that benefit societies and those that destroy them.

# The present experiment

Here we argue that a) divergent selective pressures of the social environments people are in determine behaviors that signal ‘being a good partner’ and that b) cooperation is not prosocial per se but critically depends on both the social environment and having the mere option to choose one’s social interaction partners.

*Hypothesis 1*: Norms of rule-following behaviour will establish specific to the two settings. Specifically, over time (i.e. rounds), we expect rule-following scores to decrease in the cheating setting more than in the fair setting.

*Hypothesis 2*: The first setting presented to participants will anchor the established rule-following norms in the second setting participants are exposed to. Specifically, over time (i.e. rounds), given that we expect rule-following scores to decrease in the corrupt setting, we expect rule-following scores to increase, if the dictator game was presented after.

*Hypothesis 3*: In the fair setting, those who are more rule-following will also be selected more as interaction partners.

To test these hypotheses, we conducted an experiment with a series of game-theoretic tasks in one experimental session lasting one hour.

# Ignore Everything Below

## What norms do to us

## How moral to be

## What partners to choose

Behavior that is costly to perform signals trust, a moral trait and fundamental building block of relationships.

Costly signaling theory holds that behaviors can be costly to perform but entail trust in return - regardless of whether these behaviors are intended or unintended [ref].

## When we cooperate

The effects of partner choice have mainly been investigated in cooperation settings using economic games like the Prisoner’s dilemma game, which allow for analysing the evolution of cooperation and the establishment of nomative behavior.

Indeed, effects of partner choice have mainly been investigated in cooperation settings and have been shown to … to structure social preferences, make reputation concerns more salient, …

## When we corrupt

# Methods

## Participants and ethics

Participants were recruited from the subjects pool of the Faculty of Social and Behavioural Sciences at Leiden University. Each of the 212 participants (n females; age: M = x.00 years, SD = xx.00 years) consented digitally to take part in one experimental session lasting approximately one hour and were debriefed digitally. Zero participants withdrew their participation and 12 participants were excluded due to server crashes. Therefore, the data of 200 participants was used for our analyses. Participants were paid € […] on average. Screenshots of the instructions are appended to the Supplemental Material and all data and materials are available on OSF (<https://osf.io/v4rma/>).

## Experimental design and setup

We conducted an experiment with a series of game-theoretic tasks. Figure [x] summarises both the timeline and the experimental tasks. We used one between-subjects factor (role: selector vs decider) and one within-subjects factor (setting: trust vs dishonesty). In all experimental sessions, four participants were randomly grouped together depending on their availabilities. Over a total of 30 rounds, participants went through a sequence of three stages in each round: a rule-following tasks (stage one), a partner selection task (stage two), and two settings - a trust game and a dyadic die-rolling task - for 15 rounds each (stage three). All participants were exposed to both settings and went through different stages based their role and the decisions others made. After the experimental blocks, all participants did the die-rolling task by themselves and filled in the social value orientation (SVO) scale, their demographics, and answered control questions.

* Use of oTree for programming
* Employing experimental task is a common practice both in the lab and the field [Fischbacher & Föllmi-Heusi, 2013; for a meta-analysis, see Abeler et al., 2016].

The currency in the experiment was points, where 100 points = €1.00. Participants accumulated their points individually based on their choices and over the three stages. Earnings could range from €6.50 to €14.50. Alternatively, participants could get paid out 2 SONA credits plus the bonus. If they collected less than 650 points, they were paid the show-up fee of €6.50. However, if they collected more than 650 points, all exceeding points were considered their bonus. To calculate the participants’ final payment, two rounds of the total 30 were randomly selected by the computer. The total summed amount of points they collected in these two rounds plus the points from the last die-roll task were converted to Euros and paid out to the participants. Participants earned € […] on average.

## Experimental procedure

Participants typically arrived separately at the laboratory and were seated in individual cubicles with computers. At the end of the experiment, they were asked to remain seated inside the cubicles and wait for an experimenter to open the door. We informed the participants that the present study was a group experiment in which they would interact with other participants in a group of four in total. They were also told how many of the group members already arrived and how many were still missing. We stressed this to make sure they trusted their own decisions to have an impact on other real human individuals. Participants were told that the entire experiment was computerised. The consent and debriefing forms were therefore shown on the computers. This study fell under the no-deception policy of Leiden University. So, there were neither hidden information nor deception in the study, and everything was done as stated in the instructions. After completing the comprehension checks, participants went on to the first task.

### Stage One: the rule-following task

The key variable of interest was rule-following behaviour. Participants were presented with 15 balls they could allocate one-by-one to either a blue or a yellow bucket (by clicking with the cursor on the buckets’ respective buttons on the computer screen). Participants were instructed that ‘the rule is to put the balls in the blue bucket’. They had no reason for following the rule and did not face any consequences if they did not. The rule-following variable was operationalised as the summed number of balls put in the blue bucket per round. Instructions read that each ball put in the blue bucket would earn them 5 points and that each ball put in the yellow bucket would earn them 15 points. So, choosing to put all balls in the blue bucket earned participants only a third (75 points) of what they could have earned if they put all balls in the yellow bucket (225 points). A counter below the buckets’ buttons showed the amount of money they accumulated [Kimbrough and Vostroknutov, 2015]. Participants received feedback about the scores of the other group members. To avoid reputation concerns, in all stages, feedback displays were sorted by the highest rule-following score (i.e. the summed number of balls put in the blue bucket) in descending order.

### Stage Two: partner selection

The partner selection variable was the number of partners selected per round. Participants were randomly assigned to one of two roles: selector or decider. One of the four group members was in the selector role, the other three were in the decider role. Selectors skipped the rule-following task. However, in the second stage, selectors had to select at least one of the deciders - but could also choose two or even all three deciders - for the third stage. All group members were informed that the selector had to spend 150 points for each selected decider. When selecting participants, selectors received feedback about all deciders’ rule-following scores. In the meantime, deciders had to wait. If they got selected, they had to decide about both their own and the selectors’ earnings in the third stage. Therefore, partners (i.e. selected deciders) inevitably earned more than those who were not and it was desirable to get selected. Importantly, partners interacted with the selector only, not with the other partners. Feedback about the selectors’ decisions was provided to all deciders and the feedback showed both the rule-following scores and who got chosen. We expected participants to learn from the feedback about who got chosen and who did not and adjust their choices in the next round. Participants who got chosen continued to stage three and participants who did not waited until the next round.

* we excluded the possibility to select zero deciders in order to avoid introducing powerful punishment effects [ref Fehr and gaechter?].

The partner selection variable was the number of partners selected per round. Participants were randomly assigned to one of two roles: selector or decider. Only one of the four group members was in the selector role, the other three were in the decider role. Selectors skipped the rule-following task. Instead, their role consisted on selecting at least one of the deciders - they could also choose two or all three deciders - for the third stage. Before allocating the balls to either bucket, all group members were informed that there would be a selector, who had to spend 150 points for each selected decider. When selecting participants, selectors received feedback about all deciders’ rule-following scores. In the meantime, deciders had to wait. If they got selected, they had to decide about both their own and the selectors’ earnings in the third stage. Therefore, partners (i.e. selected deciders) inevitably earned more than those who were not, and thus, it was desirable to get selected. Importantly, partners interacted with the selector only, not with the other partners. Feedback about the selectors’ decisions was provided to all deciders and the feedback showed both the rule-following scores and who got chosen. We expected participants to learn from the feedback about who got chosen and who did not and adjust their choices in the next round. Participants who got chosen as interaction partners continued to stage three and participants who did not, waited until the next round.

### Stage Three: two settings

There were two settings in the third stage: a dictator game and a dyadic die-rolling game. Each setting was repeated for 15 rounds and which one was shown first depended on the condition of the particular sessions. Therefore, the order of the settings was experimentally counterbalanced.

In one setting, participants played a modified dictator game in which selectors could not reject the offer made by the individual partners (i.e. selected deciders). Therefore, punishment could not act as a confound. In the game, the variable of interest was fairness and it was conceptualised as the amount of points that partners gave to the selectors. Partners were endowed with 500 points and had to allocate x to the selector and kept 500 - x for themselves, where x could assume a value between 0 and 250 in steps of 50 points. Options were shown in a dropdown menu. Choosing the maximum of x = 250 resulted in a fair mutual outcome and choosing the minimum of x = 0 resulted in a selfish outcome. In this stage, selectors’ earned totals were the sums of their partners’ allocated points, so that sum(1 to i) (500 - xi). Selectors received feedback about all partners’ allocations along with their rule-following scores from the first stage. Then, the next round started.

In the other setting, partners played a modified version of the dyadic die-rolling game [Weisel & Shalvi, 2015]. Instead of two participants having to match their responses, only the partners had to report a die-roll. The reported number determined partners’ own earnings and those of the selectors. The variable of interest was dishonesty and the incentives for the die-rolling game were set in such a way that participants had to be dishonest if they wanted to receive a relatively reasonable amount of points for themselves. Specifically, we interpreted dishonesty from die-rolls equal to 6 since rolling a 6 multiple times and in a row is stochastically unlikely - the probability of rolling a 6 once is 0.16, rolling a 6 twice and in a row is 0.03. Both the partners and selectors earned the same amount of points for the reported die roll, so that both earned x points. Earnings ranged from 0 points (for reporting a 1) to 250 points (for reporting a 6) in steps of 50 points. For example, if a partner reported a 4, the partner and selector both earned 150 points. However, if a partner reported a 5, both earned 200 points. At the end of every round, selectors received feedback about all partners’ contributions along with their rule-following score from the first stage. Earnings of a selector were the sum of points that each of their partners decided on or reported, so that sum(1 to i) (500 - xi). In addition, all participants received feedback about their own earnings at the end of each round. Then, the next round started. When participants reached the end of round 30, they finished the experimental block and continued to final individual-level measures.

* Employing experimental tasks such as reporting die-roll outcomes to assess dishonesty is a common practice both in the lab and
* the field (Fischbacher & Föllmi-Heusi, 2013; for a meta-analysis, see Abeler et al., 2016). At the end of each round, rule-following scores were reset to 0

### Final measures

We measured a few potential individual-level confounds. First, dishonesty could have had an influence on both the deciders’ as well as the selectors’ decisions. Similar to the previous dyadic die-roll task, this time, participants did a die-roll. However, this time, they did the die-roll for themselves only. Hence, the die-roll only increased participants’ own earnings, not the earnings of any other group members. Second, participants’ concern for others could have affected their decisions. We used a 6-item social value orientation (SVO) scale [Murphy, Ackermann & Handgraaf, 2011] slider variant with continuous choices. Last, we asked control questions about participants’ a) motives to participate, b) previous experiences with similar experiments, c) knowledge of this experiment prior to its start, and d) whether they believed they interacted with real human individuals. At the end of the experiment, participants were debriefed and informed about their final earnings digitally. They waited for an experimenter, got paid, and were thanked for their participation.

## Participants

## Material

## Procedure

## Data analysis

We used R (Version 3.5.1; R Core Team, 2018) and the R-package *papaja* (Version 0.1.0.9842; Aust & Barth, 2018) for all our analyses. Another way of citing (Aust & Barth, 2018).

# Results

# Discussion

* terms ‘selector’ and ‘decider’ confusing

# References

Abele, A. E., & Wojciszke, B. (2014). Communal and agentic content in social cognition: A dual perspective model. In *Advances in experimental social psychology* (Vol. 50, pp. 195–255). Elsevier.

Abeler, J., Becker, A., & Falk, A. (2014). Representative evidence on lying costs. *Journal of Public Economics*, *113*, 96–104.

Abeler, J., Nosenzo, D., & Raymond, C. (2016). Preferences for truth-telling.

Ades, A., & Tella, R. D. (1996). The causes and consequences of corruption: A review of recent empirical contributions. *IDs Bulletin*, *27*(2), 6–11.

André, J.-B., & Baumard, N. (2011). The evolution of fairness in a biological market. *Evolution: International Journal of Organic Evolution*, *65*(5), 1447–1456.

Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown*. Retrieved from <https://github.com/crsh/papaja>

Bahmani-Oskooee, M., & Goswami, G. G. (2005). The impact of corruption on the black market premium. *Southern Economic Journal*, 483–493.

Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution and Human Behavior*, *34*(3), 164–175.

Barclay, P., & Willer, R. (2006). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1610), 749–753.

Baumard, J., Osiurak, F., Lesourd, M., & Le Gall, D. (2014). Tool use disorders after left brain damage. *Frontiers in Psychology*, *5*, 473.

Baumard, N., André, J.-B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences*, *36*(1), 59–78.

Cialdini, R. B. (2001). Harnessing the science of persuasion. *Harvard Business Review*, *79*(9), 72–81.

Cottrell, C. A., Neuberg, S. L., & Li, N. P. (2007). What do people desire in others? A sociofunctional perspective on the importance of different valued characteristics. *Journal of Personality and Social Psychology*, *92*(2), 208.

Cukierman, A., Edwards, S., & Tabellini, G. (1989). *Seigniorage and political instability*. National Bureau of Economic Research.

Efferson, C., Roca, C. P., Vogt, S., & Helbing, D. (2016). Sustained cooperation by running away from bad behavior. *Evolution and Human Behavior*, *37*(1), 1–9.

Everett, J. A., Pizarro, D. A., & Crockett, M. (2016). Inference of trustworthiness from intuitive moral judgments. *Journal of Experimental Psychology: General*, *145*(6), 772.

Farwell, J. P. (2014). The media strategy of isis. *Survival*, *56*(6), 49–55.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*(6960), 785.

Fehr, E., & Rockenbach, B. (2004). Human altruism: Economic, neural, and evolutionary perspectives. *Current Opinion in Neurobiology*, *14*(6), 784–790.

Freud, S. (1977). *Introductory lectures on psychoanalysis*. WW Norton & Company.

Gausel, N., & Leach, C. W. (2011). Concern for self-image and social image in the management of moral failure: Rethinking shame. *European Journal of Social Psychology*, *41*(4), 468–478.

Gintis, H. (2003). Solving the puzzle of prosociality. *Rationality and Society*, *15*(2), 155–187.

Goodman, L. M. (2015). Why volkswagen cheated. *Newsweek Global*, *165*(23), 14.

Goodwin, G. P. (2015). Moral character in person perception. *Current Directions in Psychological Science*, *24*(1), 38–44.

Gupta, J. dSanjeev, Davoodi, H., & Alonso-Terme, R. (1998). *Does corruption affect income inequality and powerty?* International Monetary Fund.

Gurven, M., Allen-Arave, W., Hill, K., & Hurtado, M. (2000). “It’s a wonderful life”: Signaling generosity among the ache of paraguay. *Evolution and Human Behavior*, *21*(4), 263–282.

Hirschman, E. C. (1987). People as products: Analysis of a complex marketing exchange. *Journal of Marketing*, *51*(1), 98–108.

Hoffman, M. L. (1977). Moral internalization: Current theory and research. In *Advances in experimental social psychology* (Vol. 10, pp. 85–133). Elsevier.

Jordan, J., Mullen, E., & Murnighan, J. K. (2011). Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin*, *37*(5), 701–713.

Köbis, N. C., Prooijen, J.-W. van, Righetti, F., & Van Lange, P. A. (2016). Prospection in individual and interpersonal corruption dilemmas. *Review of General Psychology*, *20*(1), 71–85.

Lacetera, N., & Macis, M. (2010). Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme. *Journal of Economic Behavior & Organization*, *76*(2), 225–237.

Landy, J. F., Piazza, J., & Goodwin, G. P. (2016). When it’s bad to be friendly and smart: The desirability of sociability and competence depends on morality. *Personality and Social Psychology Bulletin*, *42*(9), 1272–1290.

Landy, J. F., & Uhlmann, E. L. (2018). Morality is personal. *Atlas of Moral Psychology*, 121.

Mauro, P. (1995). Corruption and growth. *The Quarterly Journal of Economics*, *110*(3), 681–712.

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

Melnikoff, D. E., & Bailey, A. H. (2018). Preferences for moral vs. Immoral traits in others are conditional. *Proceedings of the National Academy of Sciences*, *115*(4), E592–E600.

Montesquieu, C. L. (1951). Oeuvres completes (2 vols.). *Paris: Pl6iade*.

Nieto, N. (2012). Political corruption and narcotrafficking in mexico. *Transcience*, *3*(2), 24–26.

Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives*, *14*(3), 137–158.

Peeters, G. (1992). Evaluative meanings of adjectives invitro and in context-some theoretical implications and practical consequences of positive-negative asymmetry and behavioral-adaptive concepts of evaluation. *Psychologica Belgica*, *32*(2), 211–231.

Pepitone, A. (1976). Toward a normative and comparative biocultural social psychology. *Journal of Personality and Social Psychology*, *34*(4), 641.

Rand, D. G., Arbesman, S., & Christakis, N. A. (2011). Dynamic social networks promote cooperation in experiments with humans. *Proceedings of the National Academy of Sciences*, *108*(48), 19193–19198.

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>

Rose-Ackerman, S., & Palifka, B. J. (2016). *Corruption and government: Causes, consequences, and reform*. Cambridge university press.

Rothstein, B. (2011). *The quality of government: Corruption, social trust, and inequality in international perspective*. University of Chicago Press.

Sethi, R., & Somanathan, E. (1996). The evolution of social norms in common property resource use. *The American Economic Review*, 766–788.

Tomasello, M., & Rakoczy, H. (2003). What makes human cognition unique? From individual to shared to collective intentionality. *Mind & Language*, *18*(2), 121–147.

Utikal, V., & Fischbacher, U. (2013). Disadvantageous lies in individual decisions. *Journal of Economic Behavior & Organization*, *85*, 108–111.

Weisel, O., & Shalvi, S. (2015). The collaborative roots of corruption. *Proceedings of the National Academy of Sciences*, *112*(34), 10651–10656.

Wojciszke, B., Abele, A. E., & Baryla, W. (2009). Two dimensions of interpersonal attitudes: Liking depends on communion, respect depends on agency. *European Journal of Social Psychology*, *39*(6), 973–990.