



**University of
Zurich** ^{UZH}

University of Zurich
Department of Economics

Working Paper Series

ISSN 1664-7041 (print)
ISSN 1664-705X (online)

Working Paper No. 303

The Cognitive Foundations of Cooperation

Carlos Alós-Ferrer and Michele Garagnani

October 2018

The Cognitive Foundations of Cooperation*

Carlos Alós-Ferrer[†]
University of Zurich

Michele Garagnani[‡]
University of Zurich

This Version: October 2018

Abstract

Why do some individuals cooperate with their fellow human beings while others take advantage of them? The human drive for cooperation and altruism is one of the most powerful forces shaping our society, but there is an enormous behavioral variance in individual behavior. At the same time, whether it is intuitive to behave in a cooperative manner or whether such behaviors are calculated deeds remains an unanswered question. Indeed, recent empirical investigations regarding the spontaneity of human cooperation have found mixed evidence, possibly due to a failure to induce compliance in the behavioral manipulations employed. We conducted a laboratory experiment inducing intuitive and deliberative behavior through gradual economic incentives that ensure compliance. To account for individual heterogeneity, we independently measured social value orientation and aversion to interpersonal (strategic) uncertainty. We find that these measures determine the intrinsic predisposition towards cooperation. Subjects with more altruistic social values or a higher tolerance towards interpersonal uncertainty are more cooperative. Crucially, we find causal evidence that there is no universal default mode of behavior. Rather, intuition enhances intrinsic predispositions, while deliberation moderates them towards socially acceptable behavior. That is, subjects with a higher (resp. lower) predisposition towards cooperation became more (resp. less) cooperative under time pressure compared with time delay.

JEL Classification: D01 · D81 · C9

Keywords: Cooperation · Heterogeneity · Time Manipulations

Working Paper. This is an author-generated version of a research manuscript which is circulated exclusively for the purpose of facilitating scientific discussion. All rights reserved. The final version of the article might differ from this one.

*We thanks Ernst Fehr, Todd Hare, and Gustav Tinghög for discussion and comments.

[†]Corresponding author: carlos.alos-ferrer@econ.uzh.ch. Department of Economics, University of Zurich. Blümlisalpstrasse 10, 8006 Zurich, Switzerland.

[‡]Department of Economics, University of Zurich. Blümlisalpstrasse 10, 8006 Zurich, Switzerland.

Cooperation within groups has given humans a crucial evolutionary edge and remains a key factor sustaining and stabilizing human society. Accordingly, the implications of cooperative behavior have been widely investigated (Falk et al., 2003). Cooperative tendencies appear at a very young age (Warneken and Tomasello, 2006), suggesting that these characteristics are deeply rooted in our nature. And yet, self-interest often runs rampant and sometimes appears to win the upper hand in our modern society. Breakdowns of cooperative behavior threaten long-run social and economic stability, with warnings on human egoism populating scientific journals and demonstrators' banners alike.

This heterogeneity raises two fundamental questions which are crucial to understand our behavior and to inform appropriate policy interventions. The first is which are the actual determinants of cooperative behavior. The second is whether human beings are intuitively cooperative or rather instinctively predisposed towards selfishness. A few studies have focused on the first issue (Vives and FeldmanHall, 2018), and the second one has been the subject of an intense, recent debate. Empirical research has addressed this issue by exogenously manipulating the reliance on deliberation versus intuition through the available decision time, but has found generally mixed evidence (Rand et al., 2012; Rand, 2016, 2017; Tinghög et al., 2013; Bouwmeester et al., 2017; Camerer et al., 2018; Recalde et al., 2018; Gärtner, 2018). Results in favor of the social heuristic hypothesis, which declares cooperation the default mode of behavior (Rand et al., 2012; Rand, 2016, 2017), have generally not been replicated (Tinghög et al., 2013; Bouwmeester et al., 2017; Camerer et al., 2018), and the original analysis (Rand et al., 2012) has been contested (Myrseth and Wollbrant, 2017b). Some studies have tried to reconcile these mixed results by postulating that the relation between cooperation and intuition may be mediated by other factors previously neglected. For instance, evidence obtained by manipulating the location of the equilibrium in public-good games with a unique dominant strategy suggests that some decisions classified as intuitive may simply be decision errors (Recalde et al., 2018). Thus, one needs to promote as well as control for subjects' comprehension of the task. Further evidence (Gärtner, 2018) suggests that the relation between intuition and cooperation holds only if the choice is described in a cooperative frame (compared with a competitive environment). That is, the exact form in which the task is presented may cue various social norms and hence activate different heuristics.

The lack of conclusive evidence on cooperation and intuition might be closely linked to the use of excessively simple time pressure manipulations to induce intuitive behavior (Rand et al., 2012; Tinghög et al., 2013; Rand, 2016, 2017; Bouwmeester et al., 2017). In those experiments, participants are required to respond within a certain time window, the idea being that time pressure increases the reliance on impulsive processes and hence reveals the default (intuitive) mode of behavior. Conversely, other participants are forced to wait to implement their decision (time delay), to induce them to rely on more deliberative processes. However, the inference of what is intuitive and what is not may be biased by the strength of subjects' preferences. It is well-known that decision times are longer for difficult decisions, where preferences are weaker (Dashiell, 1937; Moyer and Landauer, 1967). As a consequence, both very generous and very selfish participants may respond quickly not because they are intuitively generous but rather because the decision is an easier one for them to make (Krajbich et al., 2015; Evans et al., 2015). In contrast, those with weaker preferences might miss the time window, and indeed it has been noted that many participants fail to answer within the allocated time and results often depend on whether those are included or not, to the extent that explicit calls for better manipulations have been put forward (Bouwmeester et al., 2017). Hence, the sign of the correlation between response times and contributions might simply be determined by the share of generous and selfish participants present in the experiment, and the inclusion or exclusion of those answering more slowly will change the results in arbitrary ways.

Conceptually, the debate on the cognitive determinants of cooperation is related to similar arguments with respect to a different but related dimension of human behavior, namely fairness and prosocial behavior. As in the case of cooperation, humans and other species have a highly developed sense of fairness (Brosnan and de Waal, 2014). Studies investigating the determinants of fairness suggest that there is a large heterogeneity across individuals (Fischbacher et al., 2013). Inducing intuitive behavior in dictator and ultimatum games results in different behavior in different samples, pointing at heterogeneous in the relation between intuition and fairness (Schulz et al., 2014; Achtziger et al., 2016, 2015, 2018). Some results in this literature indicated an association between short response times and fair behavior (Cappelen et al., 2016), but they have been argued to amount to a reverse inference fallacy (Myrseth and Wollbrant, 2016), and other studies have found no evidence of a relation between intuition and fair behavior (Tinghög et al., 2016). For instance, time delay or “cooling off” periods appear to have weak or no effects (Oechssler et al., 2015; Andersen et al., 2018). Recent work in the fairness literature (Chen and Krajbich, 2018) has provided evidence of a heterogeneous relation between intuition and fairness using a series of mini-dictator games where subjects were either under time pressure or time delay. Although fairness and cooperation are different aspects of human behavior, these results can be taken as additional evidence suggesting that the mixed results obtained regarding the relation of intuition and cooperation might be due to the possibly heterogeneous determinants of cooperative behavior.

In this work, we address these difficulties and take the investigation of the cognitive underpinnings of cooperative behavior in humans the a step further by proposing a new incentive-compatible framework which is immune to the strength-of-preference critique. At the same time, we identify two important sources of heterogeneity in cooperative behavior: social values and attitudes to interpersonal (strategic) uncertainty. Hence, we provide causal and conclusive evidence on the determinants of cooperative behavior. The design also addresses other critiques put forward in the recent literature, namely the appropriateness of the identification strategy (Myrseth and Wollbrant, 2016, 2017b) and the participants’ comprehension of the task (Andreoni, 1995).

Results

We collected data in 5 experimental sessions of 32 participants each, $N = 160$ (see Supplementary Materials for a power analysis). All participants made decisions in a one-shot Public Good Game (PGG), a repeated PGG, and answered questionnaires regarding external measures of pro-sociality and strategic uncertainty as well as control and demographic questions. In the PGG, participants were assigned to a group of 4 subjects. Each group member was endowed with 4 Euro and decided how much of the endowment to contribute to a common group account. The total group contributions were then doubled and split evenly among the group members. Feedback was not provided until the end of the experiment and participants had no information on the other participants they were matched with. During this task, participants were allocated to one of two possible conditions, time pressure and time delay. In the time delay condition, participants were rewarded with an additional amount of 1 Euro (on top of their earnings from the PGG) if they provided an answer after 10 seconds. In the time pressure condition, participants were endowed with an additional sum of 3 Euro, but 30 Euro cents were deducted from this amount for each second they took to make their decision (see Figure 1). This mechanism guarantees that participants under time pressure have an incentive to answer as quickly as possible, independently of how strong their preferences are (and hence how fast their response without time pressure would be). That is, both time pressure and time delay become binding to every participant independently of their preference strength.

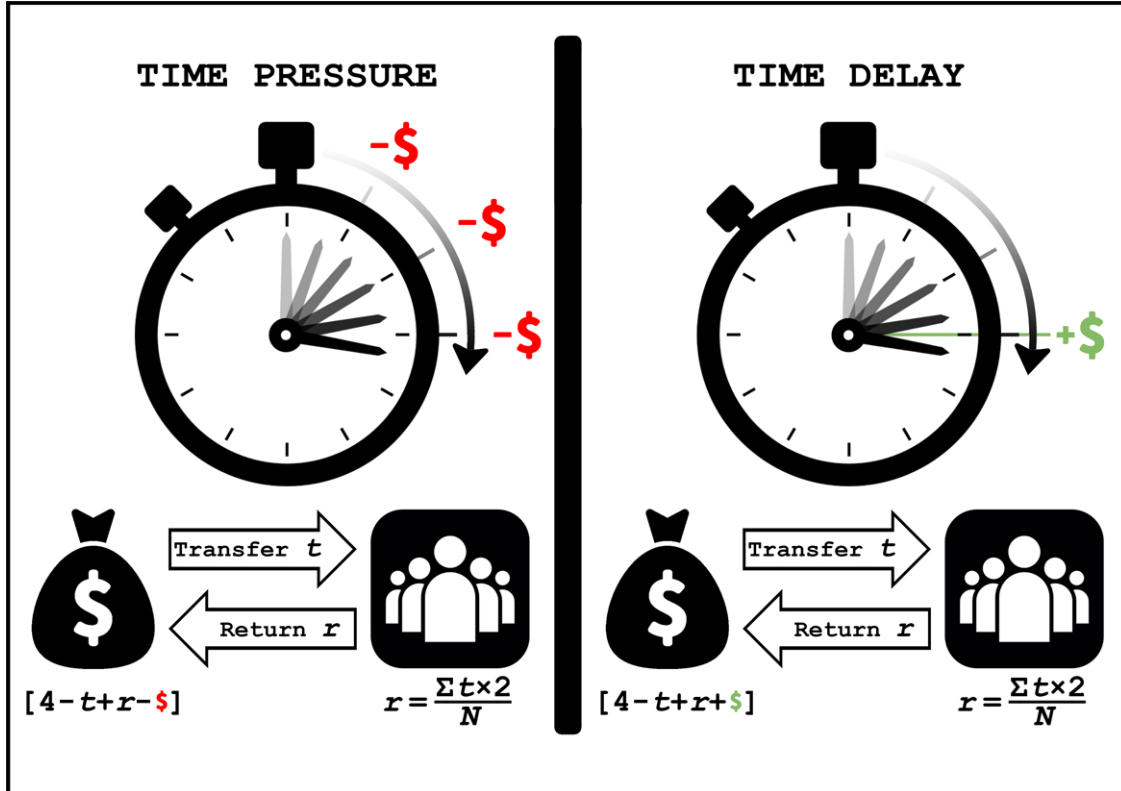


Figure 1: Participants decided how much of their endowment to contribute to a group account, and total contributions were doubled and split among the four group members. In the time delay condition (right), participants were rewarded with an additional amount of 1 Euro if they waited 10 seconds before making their choice. In the time pressure condition (left), participants were endowed with an additional sum of 3 Euro, but 30 Euro cents were detracted from this amount for each second they took to make their choice.

The majority of subjects complied with the time manipulation (139 out of 160, 86.88%) of which 78 out of 80 (97.50%) were under time delay and 61 out of 80 (76.75%) were under time pressure. We did not exclude any subjects from the analysis, but the results do not change if we exclude the subjects who did not comply with the time manipulation. To find the correct incentive level for the time manipulations, a pre-test was conducted to identify the monetary amounts guaranteeing that the earnings from the time manipulations among compliant participants did not differ across conditions (see Supplementary Materials). This was indeed the case in the main experiment (Wilcoxon rank-sum test (WRS), $N = 139$, $z = 1.352$, $P = 0.176$). Therefore, we can compare contributions across conditions to infer the differences between intuitive and deliberative choices.

The “social heuristic” effect initially reported in previous experiments is absent in our data. The fraction of the endowment contributed is not significantly different across treatments (Figure 2.A; WRS, $N = 160$, $z = 1.078$, $P = 0.281$), and the differences in the distribution of contributions across treatments are not statistically significant (Figure 2.B; Kolmogorov-Smirnov test (KS), $N = 160$, $D = 0.0625$, $P = 0.732$). Table 1 provides Tobit regressions showing that the time manipulations had no significant effect on contributions (Model 1), and that this result is robust to controlling for several possible confoundings (Rand et al., 2012; Andreoni, 1995), like understanding of the task, familiarity with the game, gender, and compliance with instructions (Model 2). The same results are obtained with linear and frac-

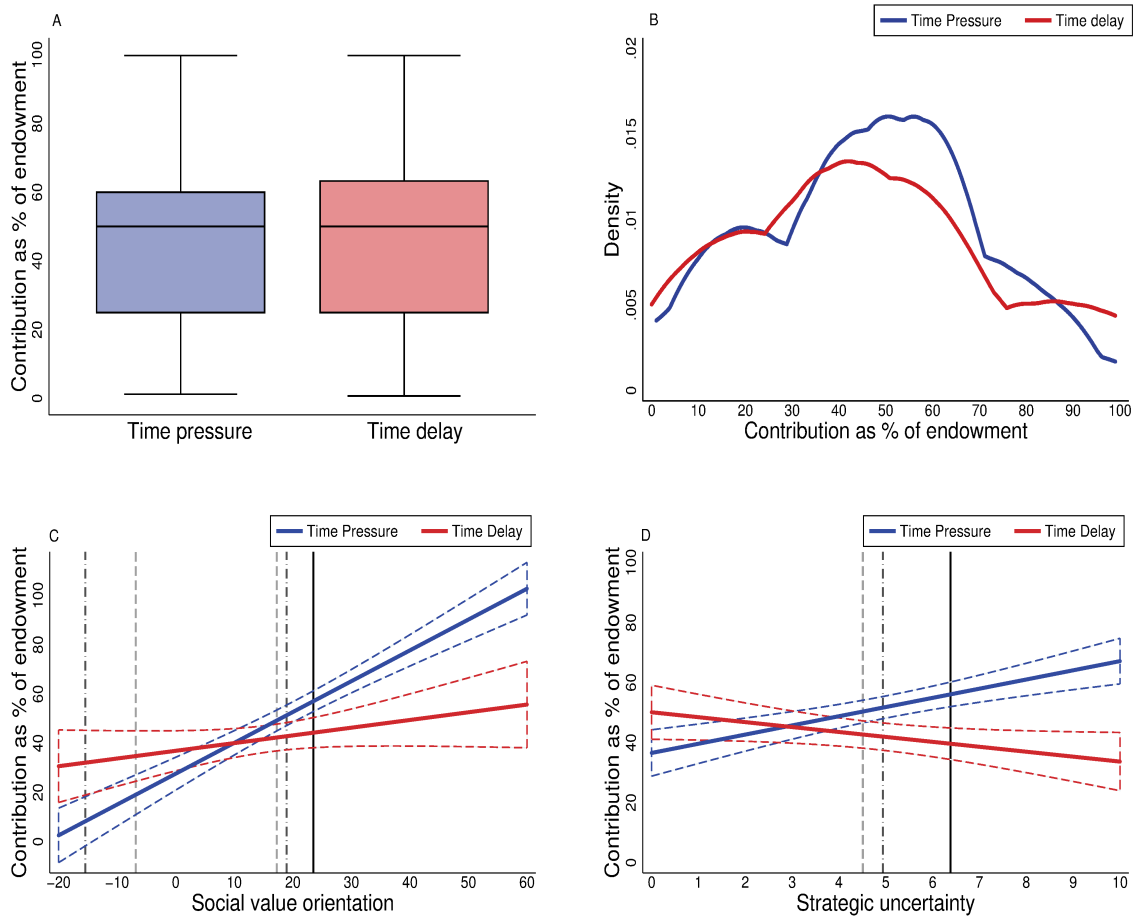


Figure 2: (A) Distribution of the contributions between treatments in the one-shot PGG, showing no treatment effect. (B) Cumulative density of the contributions between treatments in the one-shot PGG, showing no differences. (C) Plotting simple regressions of the contributions as a function of the external measure of pro-sociality shows a heterogeneous effect of the time manipulations, with more generous subjects becoming more generous under time pressure and more selfish subjects becoming even more selfish. Vertical lines indicate the Johnson-Neyman points for $P < 0.100$ (dashed grey line), $P < 0.050$ (dashed black line), and $P < 0.010$ (solid black line). Dashed colored lines indicate 95% confidence intervals. (D) A similar picture for the external measure of aversion to strategic uncertainty (SU) shows further heterogeneity, with subjects with higher (lower) SU contributing more (less) under time pressure.

tional logit regressions (Supplementary Tables 3 and 4). In summary, there is no evidence that cooperative behavior is more intuitive than selfish behavior when using incentive-compatible, binding manipulations.

These results suggest that the relation between intuition and cooperation is less straightforward than previously assumed. We claim that this relation is heavily dependent on individual heterogeneity and the determinants of cooperative behavior. To substantiate this claim, we collected two external measures of individual predispositions. The first builds upon the view that altruism and cooperation, although in principle different, are interrelated. The second takes into account that cooperation always entails a social risk.

Our first external measure is Social Value Orientation (Murphy et al., 2011) (SVO), which captures prosocial or altruistic predisposition. The score of this (incentivized) scale is summarized for each participant as an angle resulting from the weights given to outcomes for oneself and another person in a series of money allocation tasks. Roughly speaking, an angle of 0° corresponds to caring only for the own mon-

etary rewards, while the 45° angle represents giving equal weights to own and others' outcomes. A value of 22.45° is usually taken as threshold separating individualistic predispositions (below 22.45°) from more generous attitudes (above this threshold). The scale further allows to identify competitive attitudes (aiming to ensure a higher own gain compared with others) by allowing for negative values, as well as altruistic attitudes (willingness to accept penalizations in own gains to increase others' gains) with values above the diagonal. Technically, the scale takes values between -16.26° and 61.39° . Figure 2.C plots individual contributions as a function of SVO and time manipulation. Contributions are larger for larger levels of SVO in both treatments, showing that SVO is a likely determinant of cooperative behavior. The regression line becomes steeper under time pressure, indicating that more generous subjects behave more generously under time pressure but more selfish subjects behave more selfishly. Following a Johnson-Neyman (JN) procedure (Johnson and Neyman, 1936), the differences between the contributions in the two conditions are statistically significant ($P < 0.05$) at levels of pro-sociality left of -15.45° and right of 18.94° (see Supplementary Materials for details). Results are further substantiated by Models 3 and 4 of Regression Table 1 where we also control for subjects' knowledge of the PGG, experience with similar experiments, gender, compliance with the time manipulation, and understanding of the task (interaction, linear combination test, coef. = 1.532, $P < 0.001$).

The PGG entails interpersonal uncertainty, as contributing when others do not leads to the worst possible payoffs. For this reason, we also elicited attitudes towards interpersonal (strategic) uncertainty through an external (incentivized) measure (Heinemann et al., 2009) (SU). This measure is fundamentally different from individual attitudes toward risk or attitude (which are frequently measured in economics studies). Rather, it is specifically designed to capture the individual tolerance towards interpersonal uncertainty. It does so by measuring how much money a subject is willing to forgo in order to avoid a situation with interpersonal uncertainty (a coordination game) in favor of a sure but less rewarding option. We conjectured that this predisposition plays a role in determining cooperative behavior as cooperating entails the risk to be let down or taken advantage of by others. The SU scale goes from 0 to 10, with higher values indicating that the subject is more willing to accept interpersonal uncertainty to achieve potential gains and hence has a lower aversion to interpersonal strategic uncertainty (see Supplementary Materials for details). Again, we find that the external measure partially explains behavior and the relation becomes sharper under time pressure. Figure 2.D shows that subjects scoring higher (resp. lower) in SU contribute more (resp. less) under time pressure than under time delay, suggesting that time pressure makes decision makers behave more in accordance with their underlying impulses, hence contributing less for higher degrees of aversion towards strategic uncertainty. According to a JN procedure, the differences between the contributions in the two conditions are statistically significant ($P < 0.05$) at levels of strategic uncertainty bigger than 4.94. Results are substantiated by Models 5 and 6 of Regression Table 1 (interaction, linear combination test, coef. = 5.994, $P = 0.003$). Model 7 shows that the results remain when examining the impact of SVO and SU simultaneously. The same results (both for SVO and SU) are obtained with linear and fractional logit regressions (Supplementary Tables 3 and 4).

As a robustness measure (Martinsson et al., 2014; Burton-Chellew et al., 2016), we implemented a repeated PGG. In this task, participants were presented with a series of 12 one-shot PGGs, with random rematching of participants and without feedback. Figure 3.A shows a decrease in contributions over time, which is a typical finding in repeated PGGs (Rand et al., 2012), but only for time delay. Figure 3.B shows a positive relation between the average contribution in the repeated PGG and the contribution in the one-shot PGG both for time pressure (Spearman $\rho = 0.670$, $p < 0.001$) and time delay ($\rho = 0.597$, $p < 0.001$). Again, without controlling for individual preferences, there is no evidence for a relation between cooperation and intuition. The average fraction of the endowment contributed is not

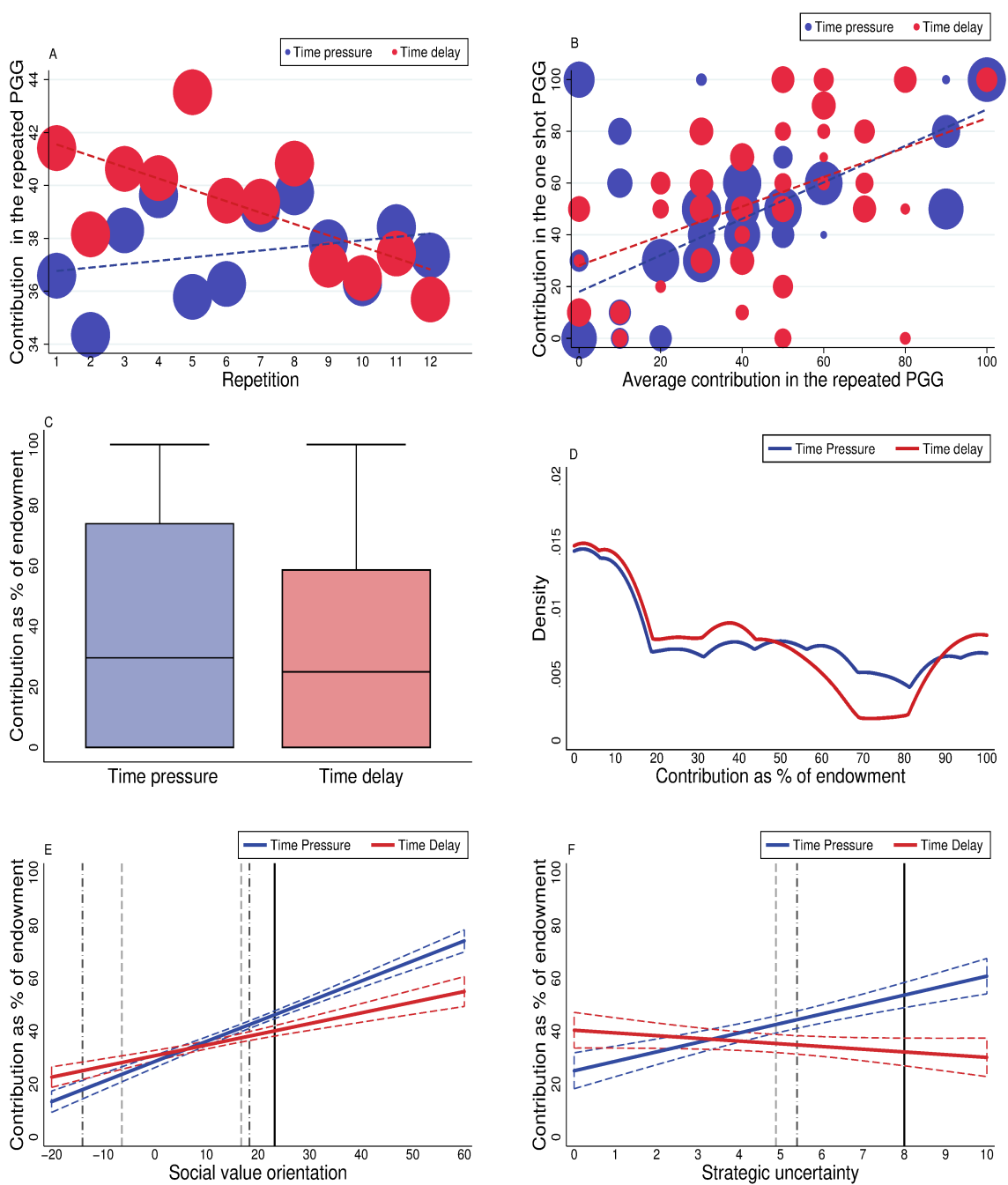


Figure 3: (A) Average contribution between treatments in each period of the repeated PGG. Dot size is proportional to the number of observations. (B) Contribution in the one-shot PGG against average contribution in the repeated PGG. (C) Distribution of contributions between treatments in the repeated PGG. (D) Cumulative density of the contribution between treatments in the repeated PGG. (E) Average contribution in the repeated PGG as a function of the external measure of pro-sociality (simple regressions). (F) Average contribution in the repeated PGG as a function of the external measure of strategic uncertainty. Johnson-Neyman procedure is reported for $P = 0.100$ (dashed grey line), $P = 0.050$ (dashed black line), and $P = 0.010$ (solid black line). Dashed colored lines indicate the 95% confidence intervals.

significantly different across treatments (Figure 3.C; WRS, $N = 160$, $z = 0.836$, $P = 0.403$), and the differences in the distribution of contributions across treatments are not statistically significant (Figure 3.D; KS, $N = 160$, $D = 0.138$, $P = 0.220$). Tobit regressions confirm that the time manipulations had no significant effect on contributions in the repeated PGG, also controlling for possible confounds (Supplementary Table 2).

Figure 3.E shows that subjects behave more generously for larger values of SVO, and that more generous subjects behave even more generously under time pressure and *vice versa* for selfish subjects. A JN procedure shows that the differences between the contributions in the two conditions are statistically

significant ($P < 0.050$) at levels of generosity left of -14.04° and right of 18.34° . The results are confirmed by Tobit regressions (Supplementary Table 2, Model 3; interaction coef. = 1.476, $P = 0.001$).

Figure 3.F shows that subjects with a higher tolerance for strategic uncertainty also contribute more on average in the PGG under time pressure than under time delay and *vice versa* for those with a lower tolerance (interaction coef. = 3.959, $P = 0.050$). A JN procedure shows that the differences between the contributions in the two conditions are statistically significant ($P < 0.050$) at levels of strategic uncertainty larger than 4.89. Again, the results are confirmed by Tobit regressions (Supplementary Table 2, Model 4)

Several studies (Burton-Chellew and West, 2013; Burton-Chellew et al., 2016; Recalde et al., 2018) have argued that repeated experiences can increase understanding of the task. The systematic relation between contributions, the external measure of pro-sociality, and tolerance toward strategic uncertainty in both the one-shot and the repeated PGG strongly suggest that the results are not due to different degrees of task comprehension (Andreoni, 1995; Ferraro and Vossler, 2010). Furthermore the results are also robust to the accumulation of evidence from previous studies. A Bayesian analysis using as prior the evidence found in comparable studies (Rand et al., 2012; Tinghög et al., 2013) shows that the data are overwhelmingly more likely to follow a model with a heterogeneous effect of the time manipulation moderated by generosity ($P(M|y) \approx 1$, $BF = 34.430$), tolerance of strategic uncertainty ($P(M|y) \approx 1$, $BF = 25.543$), or both ($P(M|y) \approx 1$, $BF = 34.846$) than the social heuristic effect ($P(M|y) < 0.001$, $BF = 7.033$; see Supplementary Materials for details).

Discussion

Cooperative behavior is a fundamental cornerstone of human society, but individual human beings are highly heterogeneous in this dimension. The determinants of this heterogeneity are largely unexplored. Further, it is also unclear whether cooperation results from intuitive human characteristics or rather from deliberative actions which overcome more selfish attitudes. In this work, we identify two key determinants explaining heterogeneity in cooperative behavior, namely social values and tolerance of interpersonal uncertainty. By proposing a properly-incentivised time manipulation to induce more deliberative versus more intuitive behavior, which answers to the call for better designs auspicated by a multi-lab registered replication report (Bouwmeester et al., 2017), and by controlling for the two determinants we identify, we provide causal evidence that intuition promotes default behavior (Pfattheicher et al., 2017; Evans et al., 2015; Chen and Krajbich, 2018) in a context of social cooperation. There is, however, a fundamental heterogeneity in what intuitive default behavior actually implies for cooperation, along the lines of arguments on the individual's strength of preference (Krajbich et al., 2015). This is the likely cause of the mixed results previously found in the literature (Rand et al., 2012; Rand, 2016, 2017; Tinghög et al., 2013; Bouwmeester et al., 2017; Camerer et al., 2018; Cappelen et al., 2016; Recalde et al., 2018; Gärtner, 2018). Investigating the cognitive determinants of behavior requires accounting for idiosyncratic characteristics, which in turn requires appropriate designs and identification strategies. These findings have further widespread repercussions, from suggesting a paradigm shift in how cognitive determinants of behavior should be investigated, to immediate policy implications (Vollan and Ostrom, 2010; Mani et al., 2013; Brozyna et al., 2018). In particular, interventions to promote cooperation can only be successful if they account for heterogeneity in behavior.

Perc. Contribution	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Time Pressure (TP)	-8.155 (8.235)	-9.425 (8.619)	-9.119 (11.560)	-15.574 (11.943)	-25.497* (15.105)	-27.271* (15.343)	-27.096* (15.705)
SVO			1.221*** (0.385)	1.018** (0.390)			0.946** (0.394)
TP × SVO			0.310 (0.558)	0.706 (0.576)			0.557 (0.585)
SU					2.299 (1.920)	2.518 (1.916)	1.528 (1.837)
TP × SU					3.695 (2.736)	3.745 (2.713)	2.790 (2.612)
<i>Controls:</i>							
Familiarity		-3.151 (3.233)		-5.022* (3.031)		-3.812 (3.155)	-5.171* (3.004)
Experience		-0.549 (0.880)		-0.119 (0.815)		-0.637 (0.860)	-0.226 (0.810)
Female		-17.040* (8.825)		-13.990* (8.288)		-12.002 (8.712)	-10.745 (8.303)
Complier		-1.523 (13.144)		3.286 (12.228)		-3.920 (12.905)	0.594 (12.224)
Understood		-5.366 (8.678)		-6.653 (8.211)		-10.045 (8.583)	-9.794 (8.251)
Constant	48.162*** (5.786)	75.424*** (20.115)	27.584*** (8.400)	56.121*** (20.522)	37.473*** (10.527)	67.810*** (21.455)	53.109** (21.324)
<i>Linear combination tests:</i>							
TP × SVO + TP			1.532*** (0.408)	1.725*** (0.412)			1.503*** 1.926
TP × SU + TP					5.994*** (1.956)	6.264*** (2.000)	4.318*** (0.419)
N	160	160	160	160	160	160	160
Pseudo R^2	0.001	0.005	0.020	0.025	0.009	0.014	0.029
Wald test	0.982	6.351	24.514***	11.853***	17.718***	36.657***	36.578***
Log likelihood	-627.394	-624.709	-615.627	-612.384	-621.958	-619.025	-609.556

Robust standard errors in brackets, * $p < 0.100$, ** $p < 0.050$, *** $p < 0.010$.

Table 1: Tobit regressions on the percentage of endowment contributed in the one-shot PGG.

Appendix: Supplementary materials

Materials and Methods

Power

Using G*Power (Faul et al., 2009), a power analysis was conducted for a Wilcoxon rank-sum test. The minimum required power for detecting a medium effect size ($d = 0.5$) was set to 0.80 (Cohen, 1992). The corresponding sample size is $N = 67$ per condition. Another power analysis was conducted for the regression analyses and for the Floodlight analyses (Spiller et al., 2013). The minimum required power for detecting a medium effect size ($f^2 = 0.09$) was set to 0.80. The corresponding sample size is $N = 45$ per condition. We obtained data from 160 individuals (80 in each condition).

Participants

Participants were recruited using ORSEE (Greiner, 2004), a standard on-line recruitment system for economic experiments, which allows for random recruitment from a predefined subject pool. Participants were German-speaking students from the University of Cologne (Germany), excluding students majoring in psychology or economics. In addition, our recruiting rules excluded participants who had previously participated in any experiment involving a Public Good Game and those participants who had previously participated in more than 30 experiments. A total of 160 participants (71 female; age range [18 – 58], average = 23.63) participated in exchange for performance-based payment plus a show-up fee of 4 Euro.

Procedure

The experiment was conducted at the Cologne Laboratory for Economic Research (CLER) using PsychoPy (Peirce, 2007) and z-Tree (Fischbacher, 2007). Experimental procedures were in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments, and also standard practices in experimental economics (e.g., no-deception rule). In agreement with the ethics and safety guidelines at the CLER, participants were all preregistered in the laboratory through ORSEE and had given written informed consent regarding the laboratory's guidelines (no further informed consent is necessary for this particular experiment). Potential participants were informed of their right to abstain from participation in the study or to withdraw consent to participate at any time without reprisal.

The experiment consisted of 5 parts: a one-shot Public Good Game (PGG), a repeated PGG, a Social Value Orientation scale (SVO)(Murphy et al., 2011), a Strategic Uncertainty scale (SU)(Heinemann et al., 2009), and a final questionnaire.

Decision times for each screen were recorded. Pay-off relevant information was fully disclosed to participants. Feedback on the earnings in each part of the experiment was provided only at the end of the experiment. A session lasted about 45 min and average earnings were 20.15 Euro (SD = 2.85) including show-up fee.

Experimental Design and Descriptive Statistics

First, participants were presented with on-screen instructions explaining the one-shot PGG and the time manipulation. The wording and the design were as close as possible to the received literature of reference (Rand et al., 2012). Participants were informed that they were assigned to a group of 4 subjects present in the same room (but were not informed about the identities of the other subjects). Each group member was

provided with a 4 Euro endowment. It was common knowledge that each member of the group would decide which percent of the endowment to allocate to a common group account, and that all the money contributed to that account would be doubled and evenly distributed among group members. Examples of the participant earnings with different allocations of endowment were provided, so as to minimize the possibility of subject confusion (Ramalingam et al., 2018). All interactions were anonymous. The decision was fully incentive compatible. Participants knew they would earn the amount not contributed to the common group account plus the amount gained from the redistribution of the common group account after the total contribution was doubled. In the one-shot PGG, individual payoffs are maximized by contributing nothing to the common group account.

Subsequently, the description of the time manipulation was presented to participants. Two types of time manipulation were implemented in the experiment, time delay and time pressure. The time manipulations were implemented between-subject. Therefore, each participant saw just one time manipulation for the entire experiment. The time manipulations were incentivised. Participants were informed that an additional source of earnings was available (other than the PGG), which would depend on the time they would take for implementing the decision. In the time-delay condition participants were paid an additional 1 Euro if they decided after 10 seconds. In the time-pressure condition participants were paid an additional 3 Euro minus 30 Euro cents for each second they took to implement their choice. In other words, they were endowed with an additional sum of 3 Euro at the beginning of the decision, but for each second they took to implement their decision 30 Euro cents were detracted from the initial amount. After 10 seconds, the additional amount of money would be 0 Euro but participants would still be able to implement their decision. Examples of the additional earnings according to different decision times were provided to participants in both conditions.

The time-pressure manipulation is therefore continuous in the time pressure condition in nature, since additional earnings decrease gradually in time, while the time-delay manipulation is discontinuous, since the additional money is available only if the threshold time is exceeded. The design was implemented in this way to avoid the misclassification of non-complying subjects under the time-delay condition. In a hypothetical continuous time-delay manipulation (where subjects would received additional money for each second they waited until the threshold was exceeded), it would not be possible to distinguish between subjects not complying with the time manipulation because they did not understand the incentive mechanism and those who valued their time more than the small increment accruing to the remaining time. Since the purpose of the manipulation is to make subjects reflect on the task at hand, it would be counter-productive to draw their attention to the evaluation of the opportunity cost of waiting. In contrast, this opportunity-cost evaluation is a key component of the time-pressure manipulation. Hence the asymmetric design is the one more adept to separately induce deliberation and intuition.

A pretest was conducted to estimate the correct payment scheme guaranteeing the equivalence between the expected earnings in both time manipulations, which then indeed held in the main experiment (WRS, $N = 139$, $z = 1.352$, $P = 0.176$). The majority of subjects complied with the time manipulation in the one shot PGG (139 out of 160, 86.88%). In the time-delay manipulation 78 of 80 (97.50%) people waited ten seconds before implementing their decision, and in the time pressure treatment 61 out of 80 (76.75%) implemented their decision within ten seconds.

Between the instructions and the decision screen, participants saw a screen asking them to press a key to start the choice phase. The “decision time” in this screen was used as a control for the heterogeneous (mechanical) swiftness of participants, average = 10.347, median = 8.301, SD = 7.666, min = 0.751, max = 33.607. Subsequently, subjects faced the decision screen, implemented as in the literature of reference (Rand et al., 2012), as a slider ranging from 0 to 4, without a predetermined starting point, and

a confirmation button. Participants used the mouse to click on the slider and then on the confirmation button to implement their choice. There was no immediate feedback after choice (only at the end of the experiment). Participants took on average 16.75 seconds to decide in the one-shot PGG, median = 15.65, SD = 10.05, min = 1.96, max = 53.78. Dividing participants by time manipulations, the average RT of compliers was 24.56 seconds under time-delay (median = 23.40, SD = 8.30, min = 11.06, max = 53.78) and 6.71 seconds under time-pressure (median = 6.94, SD = 2.15, min = 1.96, max = 9.91).

After the one-shot PGG, participants were presented with on-screen instructions for the repeated PGG, plus a repetition of the time-manipulation instructions. In the repeated PGG participants faced a series of 12 one-shot PGGs and they were told that one of them would be randomly chosen and paid. It was clearly stated that there was no link between the repeated PGG and the previous one-shot PGG. No feedback was provided between repetitions. Time manipulations were implemented as in the one-shot PGG. Therefore, in this part of the experiment participants' earnings were the sum of the non-contributed part of the endowment, the amount redistributed from the common group account, and the additional money from the time manipulation. Between the second set of instructions and the first decision screen of the repeated PGG, as well as between every round of the repetition of the PGG, participants saw a screen asking them to press a key to continue. The "decision time" of these screens is used as a control for the heterogeneous (mechanical) swiftness of participants. The decision screen for the repeated PGG was identical to the one for the one-shot PGG.

Between the second and third phase of the experiment, we implemented 4 control questions designed to ensure that subjects understood the task and their answers did not result from cognitive failure (Chou et al., 2009). In these control questions participants needed to indicate which was the contribution that would maximize group's profit, the contribution that would maximize their individual earnings, and two questions about the additional money they would earn if they would answer after 4 and 9 seconds respectively. The majority of participants understood the mechanic of the experiment, 104 (65.00%) subjects answered all control questions correctly. We added a dummy variable (Understood) to our regressions, taking value one if a participant answered all control questions correctly and zero otherwise.

After the repeated PGG, participants were informed that in the remaining parts of the experiment they would earn points which would be transformed into Euros in a proportion of 10 points for 15 Euro cents. This implementation was chosen for compatibility with previous experiments (Murphy et al., 2011; Heinemann et al., 2009). The third phase of the experiment consisted of the Social Value Orientation (SVO) scale in the 15-items slider version (Murphy et al., 2011). This scale has six primary and nine secondary items, all with the same general form. Each item is a resource-allocation choice over a well-defined continuum of joint payoffs. For example, consider a subject choosing a value x between 50 and 100 inclusive. Her payoff would be x , whereas the other's payoff would be $150 - x$. Subjects indicated their allocation choice by clicking on a slider to indicate their most preferred joint distribution. Secondary items are explicitly designed to disentangle the pro-social motivations of joint maximization from inequality aversion. One noteworthy feature of these secondary items is that all distributions correspond to perfectly equal allocations. The minimum outcome of a distribution is 15 points, while the maximum is 100 points. In this part of the experiment subjects were informed that they were randomly matched with another (unknown) participant to form a pair. One of the 15 items was randomly selected and the corresponding choice implemented, for each member of the pair. Therefore the earnings of each subject in this part of the experiment depended on the choice they made and on their choices.

Decisions in this part of the experiment can be combined into a unique score, the angle corresponding to the ratio between the average allocation to self over the average allocation to the other. Following the standard classification (Murphy et al., 2011), altruists have an angle greater than 57.15, prosocials would

have angles between 22.45 and 57.15, individualists would have angles between -12.04 and 22.45, and competitive types would have an angle less than -12.04. In our dataset, most participants were classified as Prosocial ($N=50$), Individualistic ($N=106$), and only a few as Altruistic ($N=1$) or Competitive ($N=3$). Rather than replacing a continuously measured variable with a dichotomy based on a median split of that variable, which could produce an increase in Type II errors through loss of power and an increase in Type I errors through false positives (Irwin and McClelland, 2003; McClelland et al., 2015), we implemented the external measure of pro-sociality as a continuous variable (SVO).

The SVO scale was followed by the Strategic Uncertainty (SU) scale (Heinemann et al., 2009), which indicates individual aversion to strategic risk. In this phase, the subject belongs to a group of 4 randomly selected subjects. Subjects faced 10 independent decision situations displayed in a table. In each of these 10 decision situations, the subject could choose between two options, A and B. Option A gave a sure payoff ranging from 5 to 50 points in increments of 5. The payoff for choosing option B was 50 points if at least 2 other members of the same group choose B and zero points otherwise. Typically subjects choose B for low sure payoffs of A and A for high sure payoffs. The SU value is the subject's switching point from B to A. For example, an SU of zero indicates the subject always chose the safe option, A (extremely low interpersonal risk tolerance), and an SU of ten indicates the subject always chose the interpersonally risky option, B (extremely high tolerance). It can be interpreted as the subject's certainty equivalent for strategic uncertainty in a coordination game. The higher the switching point, the less strategic-risk averse the subject is. One of the 10 decisions was selected randomly and it determined the earning of subjects for this part of the experiment. In our dataset, the average strategic uncertainty score was 4.36 (median = 4, SD = 2.84, min = 0.00, max = 10.00). The strategic uncertainty score is treated as a continuous variable, SU, following the same argument used for the measure of pro-sociality. There is a marginally significant correlation between the two external measures, with more pro-social subjects being more likely to score higher on the strategic uncertain scale (Spearman's test, $N = 160$, $\rho = 0.184$, $P = 0.063$; Pearson, $N = 160$, $\rho = 0.197$, $P = 0.046$).

The last questionnaire contained demographic questions relative to age, gender, mother tongue, experience with similar experiments, field of study, and monthly expenses. This part was not incentivized. Participants were asked about the number of experiments they had taken part in (Experience; average = 5.85, median = 5.00, SD = 5.11). Moreover, subjects reported their familiarity with PGG settings on a scale between 1 and 5 (1 = no familiarity, 5 = very familiar). The result was used to create the variable Familiarity (mean = 3.41, median = 4, SD = 1.33, min = 1, max = 5).

Robustness checks for non-parametric tests

To verify the robustness of our results, we repeat the non-parametric analysis after the exclusion of those subjects who did not comply with the time manipulations. There are no significant differences in contributions (WRS, $N = 139$, $z = 0.732$, $P = 0.464$) or their distributions (KS, $N = 139$, $D = 0.0757$, $P = 0.676$) across treatments. A similar result is obtained for the repeated PGG (WRS, $N = 139$, $z = 1.040$, $P = 0.298$; KS $N = 139$, $D = 0.1219$, $P = 0.362$).

Floodlight analysis

To clarify the interpretation of the interactions between the time manipulations and the external measure of pro-sociality, or the strategic uncertainty score, we report Floodlight analyses (Johnson and Neyman, 1936; Spiller et al., 2013) in Figure 2C-D and Figure 3 E-F. Some studies (Myrseth and Wollbrant, 2017a, 2016) propose to perform the analysis just on complying subjects, while others (Evans et al.,

2015) warn against the danger of not distinguish between subjects who understood the task and those who provide an answer derived from a cognitive failure. Therefore the Johnson-Neyman (JN) procedures are performed on the contributions of those subjects who complied with the time manipulations and successfully answered all control questions (N=104).

Using the external measure of pro-sociality as moderator of the time manipulation, in the one-shot PGG, contributions are significantly different across treatments when the external measure of pro-sociality (SVO) is outside the interval $[-6.81; 17.25]$ for $P < 0.100$ respectively, $[-15.45; 18.94]$ for $P < 0.050$ and $[-118.53; 23.50]$ for $P < 0.010$ (Figure 2.C). Similarly, in the repeated PGG, contributions are significantly different across treatments when SVO is outside the interval $[-6.39; 16.76]$ for $P < 0.100$ respectively, $[-14.04; 18.34]$ for $P < 0.050$ and $[-73.50; 23.23]$ for $P < 0.010$ (Figure 3.E). Using the strategic uncertainty score (SU) as moderator of the time manipulation, in the one-shot PGG, contributions are significantly different across treatments when SU is outside the interval $[-2.83; 4.51]$ for $P < 0.100$ respectively, $[-7.99; 4.94]$ for $P < 0.050$ and $[6.38; 19.16]$ for $P < 0.010$ (Figure 2.D). Similarly, in the repeated PGG, contributions are significantly different across treatments when SU is outside the interval $[-1.95; 4.89]$ for $P < 0.100$ respectively, $[-6.45; 5.40]$ for $P < 0.050$ and $[8.00; 13.62]$ for $P < 0.010$ (Figure 3.F).

Robustness checks, regression analysis

Table 1 in the main text contains Tobit regressions for the one-shot PGG. As robustness checks, we present alternative linear regressions in Table 3 and fractional logit regressions in Table 4, the conclusions remain unchanged. Table 5 presents alternatives random-effects linear regressions on the percentage of endowment contributed in the repeated PGG. This table has the same structure to Table 2 and the main results are unchanged.

Regression Analysis, Repeated Public Good Game

Table 2 presents random-effects Tobit regressions on the percentage contributed in the repeated PGG. Model 1 shows that there is no significant effect of the time manipulation on the amount contributed in the repeated PGG. Model 2 shows that this is robust when controlling for several possible confounding factors. Models 3 and 4 show that more generous subjects behave more generously under time pressure and *vice versa* for more selfish subjects (Tobit *post hoc* test coef. = 1.476, $P = 0.001$). Models 4 and 5 show that there is also a polarization of behavior regarding the external measure of strategic uncertainty for the repeated PGG, with participants with lower aversion to strategic uncertainty (higher SU) contributing more under time pressure (Tobit *post hoc* test, coef. = 3.959, $P = 0.050$).

Perc. Contribution	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Time Pressure	-1.711 (8.835)	-5.719 (9.344)	-0.785 (12.422)	-5.818 (12.904)	-1.970 (16.262)	-8.662 (16.537)	-5.529 (16.976)
SVO			1.310*** (0.417)	1.236*** (0.427)			1.100** (0.431)
TP × SVO			0.167 (0.597)	0.308 (0.622)			0.280 (0.633)
SU					3.889* (2.094)	4.201** (2.090)	3.091 (2.027)
TP × SU					0.069 (2.946)	0.630 (2.926)	-0.070 (2.863)
<i>Controls:</i>							
Familiarity		-1.509 (3.475)		-3.111 (3.289)		-2.146 (3.387)	-3.384 (3.253)
Experience		-0.261 (0.955)		0.143 (0.896)		-0.347 (0.929)	0.034 (0.886)
Female		-5.712 (9.515)		-2.102 (9.057)		-0.978 (9.396)	0.650 (9.065)
Complier		-18.249 (14.094)		-12.771 (13.256)		-20.127 (13.768)	-14.420 (13.174)
Understood		-6.544 (9.300)		-6.741 (8.935)		-10.932 (9.195)	-9.568 (8.931)
Constant	33.635*** (6.236)	63.611*** (21.590)	11.868 (9.119)	38.892* (22.327)	13.037 (12.533)	47.198** (23.093)	29.211 (23.216)
<i>Linear combination tests:</i>							
TP × SVO + TP			1.476*** (0.429)	1.544*** (0.437)			1.380 (2.079)
TP × SU + TP					3.959* (2.074)	4.831** (2.115)	3.021*** (0.447)
N	1920	1920	1920	1920	1920	1920	1920
Pseudo R^2	0.003	0.003	0.001	0.002	0.006	0.009	0.023
Wald test	0.037	2.579	21.723***	24.569***	7.125*	11.705	29.352***
Log likelihood	-6109.761	-6108.505	-6099.448	-6098.211	-6106.275	-6104.105	-6096.077

Standard errors in brackets, * $p < 0.100$, ** $p < 0.050$, *** $p < 0.010$.

Table 2: Random effects Tobit regressions on the percentage of endowment contributed in the repeated PGG.

Bayesian analysis

The design of the experiment was kept as close as possible to previous studies (Rand et al., 2012) to maintain comparability. Therefore we implemented similar instructions, layout of screens, control variables and incentivisation, apart from the time manipulation incentivisation. The original study (Rand et al., 2012) reported in a regression analysis, comprehending just the dummy for the treatment effect and a constant, that time pressure, compared with time delay, increased the contribution to the common group account with an average treatment effect of 10.91, $SD = 2.474$, and with a constant = 22.64, $SD = 1.524$. Assuming a normal distribution of the treatment effect, we can use this information as prior for a Bayesian normal regression (Thompson, 2014) which then estimates the posterior probability of observing a similar effect in our sample where the time pressure is properly incentivised. Compared to the frequentist analysis, there is a significant positive effect of the treatment alone (coef. = 11.717, 95% Cred.Interval: [8.960; 14.697]), indicating that when the evidence is cumulated from the original study

there is evidence for generosity being intuitive. However, if we assume a more agnostic position, by using an uninformed prior for our analysis or the results obtained in other studies about this topic (5), we observe no significant treatment effect of time pressure on the amount contributed to the one shot PGG (coef. = -4.665, 95% Cred.Interval: [-15.036; 5.772]). These results indicate that the evidence in favour of generosity being intuitive in our dataset is not robust. Moreover a model comparison (Kass and Raftery, 1995; Berger, 2006) shows that even when using the informed prior from the original study (Rand et al., 2012), the model comprehending just the treatment variable has a posterior probability ($P(M|y) < 10^{-5}$ and $BF = 7.033$ (Jeffreys, 1961)) compared with the model which also comprehends the interaction between the external measure of pro-sociality and its interaction with the time pressure ($P(M|y) \approx 1.000$ and $BF = 34.430$). When we introduce into the analysis the external measure of strategic uncertainty, we can see that this variable is less relevant than the external measure of pro-sociality. A comparison, using the informed prior, between the model with the interaction between the time manipulation and the measure of pro-sociality ($P(M|y) = 0.996$ and $BF = 34.430$) as compared with a model with the interaction between the time manipulation and the measure of strategic risk ($P(M|y) < 0.001$ and $BF = 25.543$) shows that our data are less likely to follow the last model. However both models are outperformed by the combined model, which comprehends both external measures and their respective interactions with the time manipulation ($P(M|y) \approx 1.000$, $BF = 34.846$). The Bayesian analysis indicates that our data are more likely following a model where there is an heterogeneous effect of the time manipulations moderated by both generosity and strategic risk propensity.

Supplementary tables

Perc. Contribution	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Time Pressure	-4.525 (5.637)	-5.319 (6.126)	-2.921 (7.718)	-7.064 (7.797)	-13.552 (9.634)	-14.416 (9.836)	-12.353 (10.107)
SVO			0.929*** (0.231)	0.789*** (0.225)			0.747*** (0.226)
TP × SVO			0.048 (0.340)	0.318 (0.351)			0.227 (0.357)
SU					1.646 (1.265)	1.707 (1.253)	0.978 (1.231)
TP × SU					1.940 (1.812)	1.920 (1.789)	1.347 (1.784)
<i>Controls:</i>							
Familiarity		-1.589 (2.134)		-2.843 (2.024)		-1.927 (2.094)	-2.892 (1.995)
Experience		-0.356 (0.639)		-0.105 (0.594)		-0.404 (0.626)	-0.155 (0.583)
Female		-12.745** (6.076)		-10.439* (5.617)		-9.829* (5.870)	-8.689 (5.493)
Complier		-0.718 (9.902)		3.004 (8.745)		-2.661 (8.913)	1.333 (8.303)
Understood		-3.905 (6.666)		-4.353 (6.444)		-6.845 (6.458)	-5.968 (6.464)
Constant	47.938*** (3.756)	64.917*** (15.222)	32.479*** (5.799)	49.840*** (14.664)	40.199*** (6.564)	59.507*** (14.509)	47.546*** (14.581)
<i>Linear combination tests:</i>							
TP × SVO + TP			0.977*** (0.250)	1.106*** (0.260)			0.974* (1.292)
TP × SU + TP					3.586*** (1.297)	3.627*** (1.298)	2.325*** (0.260)
N	160	160	160	160	160	160	160
R^2 adjusted	-0.002	0.002	0.127	0.133	0.040	0.042	0.141
F statistics	0.644	1.070	11.557***	6.550***	3.352***	2.383***	6.025***
Log likelihood	-797.834	-794.956	-785.789	-782.640	-793.390	-790.595	-780.793

Robust standard errors in brackets, * $p < 0.100$, ** $p < 0.050$, *** $p < 0.010$.

Table 3: Linear regression on the percentage of endowment contributed in the one shot PGG.

Prop. Contribution	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Time Pressure	-0.184 (0.227)	-0.220 (0.248)	-0.139 (0.357)	-0.318 (0.359)	-0.584 (0.418)	-0.625 (0.423)	-0.570 (0.474)
SVO			0.040*** (0.011)	0.035*** (0.010)			0.033*** (0.010)
TP × SVO			0.002 (0.016)	0.014 (0.017)			0.010 (0.016)
SU					0.067 (0.051)	0.072 (0.051)	0.044 (0.052)
TP × SU					0.083 (0.076)	0.083 (0.075)	0.061 (0.078)
<i>Controls:</i>							
Familiarity		-0.065 (0.086)		-0.126 (0.087)		-0.081 (0.086)	-0.130 (0.086)
Experience		-0.015 (0.026)		-0.004 (0.026)		-0.018 (0.026)	-0.007 (0.025)
Female		-0.520** (0.245)		-0.458* (0.238)		-0.411* (0.240)	-0.385 (0.234)
Complier		-0.028 (0.407)		0.131 (0.393)		-0.097 (0.382)	0.064 (0.385)
Understood		-0.163 (0.270)		-0.176 (0.275)		-0.295 (0.270)	-0.256 (0.281)
Constant	-0.074 (0.150)	0.619 (0.621)	-0.734*** (0.259)	0.007 (0.630)	-0.391 (0.266)	0.400 (0.600)	-0.087 (0.636)
<i>Linear combination tests:</i>							
TP × SVO + TP			0.042*** (0.012)	0.048*** (0.012)			0.043* (0.058)
TP × SU + TP					0.150*** (0.057)	0.155*** (0.056)	0.105*** (0.012)
N	160	160	160	160	160	160	160
Pseudo R^2	0.002	0.015	0.054	0.068	0.022	0.034	0.075
Wald test	0.658	6.308	27.053***	38.999***	9.212***	16.415***	43.120***
Log likelihood	-110.185	-108.737	-104.346	-102.900	-107.936	-106.558	-102.044

Robust standard errors in brackets, * $p < 0.100$, ** $p < 0.050$, *** $p < 0.010$.

Table 4: Fractional logit regression on the proportion of endowment contributed in the one shot PGG.

Perc. Contribution	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Time Pressure	-1.702 (5.034)	-4.869 (5.395)	0.230 (7.102)	-3.922 (7.504)	-1.986 (9.343)	-6.735 (9.654)	-4.086 (9.963)
SVO			0.791*** (0.240)	0.732*** (0.250)			0.662*** (0.255)
TP × SVO			0.001 (0.342)	0.105 (0.362)			0.088 (0.373)
SU					2.100* (1.200)	2.213* (1.217)	1.564 (1.194)
TP × SU					0.072 (1.688)	0.394 (1.703)	0.035 (1.687)
<i>Controls:</i>							
Familiarity		-1.259 (2.016)		-2.194 (1.929)		-1.604 (1.985)	-2.329 (1.925)
Experience		-0.128 (0.544)		0.081 (0.516)		-0.192 (0.535)	0.014 (0.515)
Female		-6.428 (5.497)		-4.160 (5.286)		-3.938 (5.476)	-2.762 (5.332)
Complier		-13.622* (8.185)		-10.692 (7.766)		-14.817* (8.079)	-11.682 (7.794)
Understood		-3.368 (5.346)		-3.205 (5.198)		-5.654 (5.332)	-4.638 (5.236)
Constant	39.353*** (3.560)	62.567*** (12.528)	26.194*** (5.222)	48.179*** (13.070)	29.483*** (6.644)	54.209*** (13.497)	43.477*** (13.677)
<i>Linear combination tests:</i>							
TP × SVO + TP			0.792*** (0.243)	0.837*** (0.252)			0.750 (1.224)
TP × SU + TP					2.172*** (1.187)	2.607*** (1.228)	1.600*** (0.261)
N	1920	1920	1920	1920	1920	1920	1920
R ² overall	0.001	0.018	0.068	0.088	0.030	0.054	0.119
Wald test	0.114	4.265	21.565***	25.210***	6.526*	12.014	28.773***

Standard errors in brackets, * $p < 0.100$, ** $p < 0.050$, *** $p < 0.010$.

Table 5: Random effects regressions on percentage of endowment contributed in the repeated PGG.

References

- Achtziger, A., Alós-Ferrer, C., and Wagner, A. K. (2015). Money, Depletion, and Prosociality in the Dictator Game. *Journal of Neuroscience, Psychology, and Economics*, 8(1):1–14.
- Achtziger, A., Alós-Ferrer, C., and Wagner, A. K. (2016). The Impact of Self-Control Depletion on Social Preferences in the Ultimatum Game. *Journal of Economic Psychology*, 53:1–16.
- Achtziger, A., Alós-Ferrer, C., and Wagner, A. K. (2018). Social Preferences and Self-Control. *Journal of Behavioral and Experimental Economics*, forthcoming.
- Andersen, S., Gneezy, U., Kajackaite, A., and Marx, J. (2018). Allowing for Reflection Time Does not Change Behavior in Dictator and Cheating Games. *Journal of Economic Behavior & Organization*, 145:24–33.
- Andreoni, J. (1995). Cooperation in Public-Goods Experiments: Kindness or Confusion? *American Economic Review*, pages 891–904.
- Berger, J. (2006). The Case for Objective Bayesian Analysis. *Bayesian Analysis*, 1(3):385–402.
- Bouwmeester, S., Verkoeijen, P. P., Aczel, B., Barbosa, F., Bègue, L., Brañas-Garza, P., Chmura, T. G., Cornelissen, G., Døssing, F. S., Espín, A. M., et al. (2017). Registered Replication Report: Rand, Greene, and Nowak (2012). *Perspectives on Psychological Science*, 12(3):527–542.
- Brosnan, S. F. and de Waal, F. B. (2014). Evolution of Responses to (Un)Fairness. *Science*, 346(6207):1251776.
- Brozyna, C., Guilfoos, T., and Atlas, S. (2018). Slow and Deliberate Cooperation in the Commons. *Nature Sustainability*, 1(4):184.
- Burton-Chellew, M. N., El Mouden, C., and West, S. A. (2016). Conditional Cooperation and Confusion in Public-Goods Experiments. *Proceedings of the National Academy of Sciences*, 113(5):1291–1296.
- Burton-Chellew, M. N. and West, S. A. (2013). Prosocial Preferences do not Explain Human Cooperation in Public-Goods Games. *Proceedings of the National Academy of Sciences*, 110(1):216–221.
- Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmeld, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., Isaksson, S., Manfredi, D., Rose, J., Wagenmakers, E.-J., and Wu, H. (2018). Evaluating the Replicability of Social Science Experiments in Nature and Science Between 2010 and 2015. *Nature Human Behaviour*, 2(9):637–644.
- Cappelen, A. W., Nielsen, U. H., Tungodden, B., Tyran, J.-R., and Wengström, E. (2016). Fairness is Intuitive. *Experimental Economics*, 19(4):727–740.
- Chen, F. and Krajbich, I. (2018). Biased Sequential Sampling Underlies the Effects of Time Pressure and Delay in Social Decision Making. *Nature Communications*, 9(3557):1–10.
- Chou, E., McConnell, M., Nagel, R., and Plott, C. R. (2009). The Control of Game form Recognition in Experiments: Understanding Dominant Strategy Failures in a Simple two Person Guessing Game. *Experimental Economics*, 12(2):159–179.
- Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, 112(1):155.
- Dashiell, J. F. (1937). Affective Value-Distances as a Determinant of Aesthetic Judgment-Times. *American Journal of Psychology*, 50:57–67.
- Evans, A. M., Rand, D. G., and Dillon, K. D. (2015). Fast But Not Intuitive, Slow But Not Reflective: Decision Conflict Drives Reaction Times in Social Dilemmas. *Journal of Experimental Psychology: General*, 144(5):951–966.
- Falk, A., Fehr, E., and Fischbacher, U. (2003). On the Nature of Fair Behavior. *Economic Inquiry*, 41(1):20–26.

- Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical Power Analyses Using G* Power 3.1: Tests for Correlation and Regression Analyses. *Behavior Research Methods*, 41(4):1149–1160.
- Ferraro, P. J. and Vossler, C. A. (2010). The Source and Significance of Confusion in Public Goods Experiments. *The BE Journal of Economic Analysis & Policy*, 10(1).
- Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. *Experimental Economics*, 10(2):171–178.
- Fischbacher, U., Hertwig, R., and Bruhin, A. (2013). How to Model Heterogeneity in Costly Punishment: Insights from Responders' Response Times. *Journal of Behavioral Decision Making*, 26(5):462–476.
- Gärtner, M. (2018). The Prosociality of Intuitive Decisions Depends on the Status Quo. *Journal of Behavioral and Experimental Economics*, 74:127–138.
- Greiner, B. (2004). An Online Recruiting System for Economic Experiments. *Forschung und wissenschaftliches Rechnen GWDG Bericht*, 63(2):79–93.
- Heinemann, F., Nagel, R., and Ockenfels, P. (2009). Measuring Strategic Uncertainty in Coordination Games. *Review of Economic Studies*, 76(1):181–221.
- Irwin, J. R. and McClelland, G. H. (2003). Negative Consequences of Dichotomizing Continuous Predictor Variables. *Journal of Marketing Research*, 40(3):366–371.
- Jeffreys, H. (1961). *The Theory of Probability*. OUP Oxford, Oxford.
- Johnson, P. O. and Neyman, J. (1936). Tests of Certain Linear Hypotheses and Their Application to Some Educational Problems. *Statistical Research Memoirs*, 1(1):57–93.
- Kass, R. E. and Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90(430):773–795.
- Krajbich, I., Bartling, B., Hare, T., and Fehr, E. (2015). Rethinking Fast and Slow Based on a Critique of Reaction-time Reverse Inference. *Nature Communications*, 6:7455.
- Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, 341(6149):976–980.
- Martinsson, P., Myrseth, K. O. R., and Wollbrant, C. (2014). Social Dilemmas: When Self-Control Benefits Cooperation. *Journal of Economic Psychology*, 45:213–236.
- McClelland, G. H., Lynch, J. G., Irwin, J. R., Spiller, S. A., and Fitzsimons, G. J. (2015). Median Splits, Type II Errors, and False-Positive Consumer Psychology: Don't Fight the Power. *Journal of Consumer Psychology*, 4(25):679–689.
- Moyer, R. S. and Landauer, T. K. (1967). Time Required for Judgements of Numerical Inequality. *Nature*, 215(5109):1519–1520.
- Murphy, R. O., Ackermann, K. A., and Handgraaf, M. J. J. (2011). Measuring Social Value Orientation. *Judgment and Decision Making*, 6(8):771–781.
- Myrseth, K. O. R. and Wollbrant, C. E. (2016). Commentary: Fairness is Intuitive. *Frontiers in Psychology*, 7(654).
- Myrseth, K. O. R. and Wollbrant, C. E. (2017a). Cognitive Foundations of Cooperation Revisited: Commentary on Rand et al. (2012, 2014). *Journal of Behavioral and Experimental Economics*, forthcoming.
- Myrseth, K. O. R. and Wollbrant, C. E. (2017b). Cognitive Foundations of Cooperation Revisited: Commentary on Rand et al.(2012, 2014). *Journal of Behavioral and Experimental Economics*, 69:133–138.
- Oechssler, J., Roider, A., and Schmitz, P. W. (2015). Cooling Off in Negotiations: Does it Work? *Journal of Institutional and Theoretical Economics JITE*, 171(4):565–588.

- Peirce, J. W. (2007). PsychoPy Psychophysics Software in Python. *Journal of Neuroscience Methods*, 162(1):8–13.
- Pfattheicher, S., Keller, J., and Knezevic, G. (2017). Sadism, the intuitive system, and antisocial punishment in the public goods game. *Personality and Social Psychology Bulletin*, 43(3):337–346.
- Ramalingam, A., Morales, A. J., and Walker, J. M. (2018). Varying Experimental Instructions to Improve Comprehension: Punishment in Public Goods Games. *Journal of Behavioral and Experimental Economics*, 73:66–73.
- Rand, D. G. (2016). Cooperation, Fast and Slow: Meta-Analytic Evidence for a Theory of Social Heuristics and Self-Interested Deliberation. *Psychological Science*, 27(9):1192–1206.
- Rand, D. G. (2017). Reflections on the Time-Pressure Cooperation Registered Replication Report. *Perspectives on Psychological Science*, 12(3):543–547.
- Rand, D. G., Greene, J. D., and Nowak, M. A. (2012). Spontaneous Giving and Calculated Greed. *Nature*, 489:427–430.
- Recalde, M. P., Riedl, A., and Vesterlund, L. (2018). Error-prone inference from response time: The case of intuitive generosity in public-good games. *Journal of Public Economics*, 160:132147.
- Schulz, J. F., Fischbacher, U., Thöni, C., and Utikal, V. (2014). Affect and Fairness: Dictator Games under Cognitive Load. *Journal of Economic Psychology*, 41:77–87.
- Spiller, S. A., Fitzsimons, G. J., Lynch Jr, J. G., and McClelland, G. H. (2013). Spotlights, Floodlights, and the Magic Number Zero: Simple Effects Tests in Moderated Regression. *Journal of Marketing Research*, 50(2):277–288.
- Thompson, J. (2014). *Bayesian Analysis with Stata*. Stata Press, Lakeway Drive, USA.
- Tinghög, G., Andersson, D., Bonn, C., Böttiger, H., Josephson, C., Lundgren, G., Västfjäll, D., Kirchler, M., and Johannesson, M. (2013). Intuition and Cooperation Reconsidered. *Nature*, 498(7452):E1–E2.
- Tinghög, G., Andersson, D., Bonn, C., Johannesson, M., Kirchler, M., Koppel, L., and Västfjäll, D. (2016). Intuition and Moral Decision-Making—the Effect of Time Pressure and Cognitive Load on Moral Judgment and Altruistic Behavior. *PloS One*, 11(10):e0164012.
- Vives, M.-L. and FeldmanHall, O. (2018). Tolerance to Ambiguous Uncertainty Predicts Prosocial Behavior. *Nature Communications*, 9(1):2156–2165.
- Vollan, B. and Ostrom, E. (2010). Cooperation and the Commons. *Science*, 330(6006):923–924.
- Warneken, F. and Tomasello, M. (2006). Altruistic Helping in Human Infants and Young Chimpanzees. *Science*, 311(5765):1301–1303.