Cooperation in public good games. Calculated or confused?

Goeschl, Timo; Lohse, Johannes

DOI:
10.1016/j.euroecorev.2018.05.007

Cooperation in Public Good Games. Calculated or Confused?

Timo Goeschl, Johannes Lohse

PII: S0014-2921(18)30085-0
DOI: 10.1016/j.euroecorev.2018.05.007
Reference: EER 3151

To appear in: *European Economic Review*

Received date: 10 May 2017
Accepted date: 28 May 2018

Please cite this article as: Timo Goeschl, Johannes Lohse, Cooperation in Public Good Games. Calculated or Confused?, *European Economic Review* (2018), doi: 10.1016/j.euroecorev.2018.05.007

Cooperation in Public Good Games. Calculated or Confused?

Timo Goeschl[a]
University of Heidelberg

Johannes Lohse[b]
University of Birmingham

June 12, 2018

Abstract

Some recent experimental papers have claimed that contribution decisions in a public goods game (PGG) are more likely to be cooperative if based on intuition rather than reflection. In light of conflicting findings, this paper (i) reinvestigates the behavioral impact of so-called cognitive style in the PGG; and (ii) connects it with an earlier literature on the role of cognitive failure (confusion). This is motivated by the possibility that the method of time pressure, commonly used to identify cognitive style, invites confusion as a confounding factor. Two channels for such confounds are identified and experimentally tested: A heterogeneous treatment effect of time pressure depending on subjects' confusion status and a direct impact of time pressure on subjects' likelihood of being confused. Our reinvestigation of the behavioral impact of time pressure confirms that cognitive style matters, but that deliberation rather than intuition drives cooperation. The confounding effect of confusion is not found to be direct, but to operate through a heterogeneous treatment effect. Time pressure selectively reduces average contributions among those subjects whose contributions can confidently be interpreted as cooperative rather than confused.

[a]Email: goeschl@eco.uni-heidelberg.de. Postal address: Department of Economics, Bergheimer Str. 20, 69115 Heidelberg, Germany.

[b]Corresponding Author. Email: J.Lohse@bham.ac.uk Postal address: JG Smith Building, University of Birmingham, Edgbaston Campus, B152SB Birmingham

1 Introduction

Over the last ten years, economists have increasingly embraced so-called 'dual-system' models pioneered in psychology (Loewenstein and O' Donoghue, 2007; Evans, 2008; Kahneman, 2011; Dreber et al., 2014). Such models distinguish between intuition and deliberation as the two main "cognitive styles", that is, the main types of mental processes used in decision-making. One area of particular interest has been whether this distinction of cognitive styles can help to advance our still incomplete, and sometimes conflicting (Dreber et al., 2014), understanding of cooperation in social dilemmas, asking - in a nutshell - if cognitive style influences cooperation or pro-social behavior more generally (Piovesan and Wengström, 2009; Fiedler et al., 2013b; Ubeda, 2014; Grossman et al., 2014; Martinsson et al., 2014; Corgnet et al., 2015; Achtziger et al., 2015; Hauge et al., 2016; Kocher et al., 2017; Lohse et al., 2017; Merkel and Lohse, 2018). Some papers claim to have answered this question in the affirmative: Rand et al. (2012) and Rand et al. (2014) suggest that a positive link can be established between subjects using an intuitive cognitive style in social dilemma situations and their propensity to act cooperatively; and they suggest that this link is causal.

A key challenge for exploring the link between cognitive style and cooperation is that the cognitive style used in a given decision situation is not directly observable. When analyzing choice data from social dilemma experiments, researchers therefore need to infer the cognitive style used by other means. In observational studies, they exploit additional non-choice data on subjects' response time, relying on the observation that fast decisions are on average more likely to be associated with a more intuitive cognitive style (Rubinstein, 2007; Schotter, 2008; Schotter and Trevino, 2014; Spiliopoulos and Ortmann, 2016). Studies aiming at establishing causality exogenously restrict the time available for making a decision in order to manipulate the speed with which experimental subjects decide about cooperation, thereby affecting the use of deliberation in the choice process. Such "time pressure" treatments now constitute a popular methodology in many experiments in which cognitive style is an object of inquiry, including the studies by Rand et al. (2012) and Rand et al. (2014). A judicious interpretation is advised, however: Several papers have highlighted how evidence from time-pressure manipulations or response time correlations can mislead the researcher (Tinghög et al., 2013; Krajbich et al., 2015; Bouwmeester et al., 2017; Myrseth and Wollbrant, 2017; Merkel and Lohse, 2018).[1]

The original claim of "spontaneous giving and calculated greed" (Rand et al., 2012) has attracted considerable interest, with much of the follow-up work focusing on the robustness of the evidence that time pressure causes contributions in the PGG to increase. Cone and Rand (2014) and Rand et al. (2015) report additional supportive evidence and a subsequent meta analysis of 51 studies of social dilemma experiments that manipulate cognitive processing by time pressure or

[1]A clean interpretation can be complicated due to selection effects between compliers and non-compliers (Tinghög et al., 2013; Bouwmeester et al., 2017), differences in the strength of preferences (Krajbich et al., 2015; Merkel and Lohse, 2018), or due to the fact that despite a binding time limit average decision times might exceed thresholds conventionally associated with intuitive decision making in cognitive psychology (Myrseth and Wollbrant, 2017). For instance, re-analysing the original data in Rand et al. (2012) with these concerns in mind suggests that the original effect is only weakly significant when controlling for selection effects (Tinghög et al., 2013) and is not present at all when focussing on a subset of very fast decisions, for which it is most plausible that subjects have relied on intuition (Myrseth and Wollbrant, 2017).

other means finds a significant positive average effect of intuition on contributions[2], but also a high degree of variability across the studies included (Rand, 2016). Other studies such as Tinghög et al. (2013) and Verkoeijen and Bouwmeester (2014), however, have failed to replicate the original findings. Using a different design, Duffy and Smith (2014) and Martinsson et al. (2014) also find no evidence for intuitive cooperation in repeated public good games. The most rigorous reinvestigation is a replication exercise (Bouwmeester et al., 2017) that combines data from twenty-one separate pre-registered replications of the causal experiment reported in Rand et al. (2012). The results of this exercise not only underline the earlier finding of high variability across studies. It also suggests that selection effects best explain the original evidence and that there is little support for the original notion that time pressure causes higher contribution levels in the PGG.

The present paper, while also providing additional evidence on the effect of time pressure on contribution levels in the PGG, mainly focuses on a part of the original claim that has attracted less attention. Even if time pressure led to changes in contribution levels in the PGG, could such changes be validly interpreted as changes in cooperation? An economics literature that precedes the discussion on a link between *cognitive style* and cooperation has demonstrated that there is a link between *cognitive failure* and observed behavior in the PGG. In a seminal paper, Andreoni (1995) showed that a significant share of behavior in the PGG that appeared to be cooperation could in fact be traced back to confusion about core elements of the strategic situation.[3] These findings raise the possibility that any link between time pressure and higher contributions cannot only be attributed to differences in cognitive style (through decision speed), but could also be due to cognitive failure (through confusion). Cognitive failure could conceivably confound the inference from time pressure experiments through two channels. The first channel is a heterogeneous treatment effect of time pressure on subjects, depending on whether they are confused or unconfused about material aspects of the choice situation at the exact moment when they are asked to select their preferred contribution level.[4] For instance, some subjects might simply have misread the instructions *before* they reach the decision screen and hence do not understand that there is a trade-off between selfish and other-regarding choices. If subjects respond differently to time pressure depending on their confusion status, then the presence and sign of the average treatment effect across all subjects, confused and unconfused, could be misleading. Evidence in favour of selective treatment effects of this kind is found in Strømland et al. (2016). In the presence of heterogeneous treatment effects, different average treatment effects found across single experiments could well depend on the (unobserved) level of confusion in each study population, reflecting, for example, demographic composition or the format of the

[2]The effect size is 4.2 percentage points in the full sample and 6.1 percentage points in a restricted sample of compliers.

[3]Subsequent research has been able to confirm the important role of confusion for explaining positive contributions in the PGG and has demonstrated that repeated experiences and learning can attenuate its impact over time (Houser and Kurzban, 2002; Ferraro and Vossler, 2010; Burton-Chellew and West, 2013; Bayer et al., 2013; Burton-Chellew et al., 2016).

[4]From here onwards, we will refer to confusion as a selfish subject's cognitive failure to understand that free-riding is his pay-off maximizing strategy. More generally, Chou et al. (2009, p.160) use the term *game form recognition* to specify different kinds of cognitive failure. Subjects can be said to display a perfect game form recognition (i.e. they are unconfused), if they understand "(1) the sets of strategies available [...], (2) the information conditions, (3) the relationship between strategy choices and outcomes, and (4) the relationships between outcomes and payoffs."

instructions (Ferraro and Vossler, 2010).

The second conceivable confounding channel is the possibility that subjecting participants in experiments to time pressure affects their likelihood of cognitive failure. Taxing subjects' deliberative capacities by applying time pressure could plausibly lead to a higher share of confused subjects. Recent correlational evidence indicates that this possibility exists: in Rubinstein (2013), those subjects who decided faster and were therefore more likely to make an intuitive choice, were also significantly more likely to choose strictly dominated actions, a clear indicator of confusion. Similarly, a response time study of Recalde et al. (2014) finds that faster participants in a non-linear PGG were more likely to contribute at suboptimal levels given their own preferences. Faster decisions could therefore systematically correlate with the likelihood that the decider is confused about core features of the decision. While the existing evidence is merely correlational at this stage, it opens up the possibility that time pressure could in fact cause participants to become more confused. We conduct a direct test for this causal relationship in the current paper. If confirmed, higher average contributions under time pressure - as found in Rand et al. (2012) - could no longer be equated with more cooperation and hence could shed no light on a link between cognitive style and cooperation.

Against this background, the present paper pursues two objectives. The first is to reinvestigate the causal link between cognitive style and cooperation and to learn more about its robustness. The goal is to emulate earlier experiments that combine a linear PGG with a time pressure manipulation in order to reproduce the treatment effect of exogenously induced intuition on behavior. This requires a baseline implementation that is designed to maximize comparability with the published literature. On robustness, we explore two dimensions. The first dimension answers to a noted source of complications for interpreting the evidence from existing time pressure studies, namely that subjects do not consistently obey the time constraint (Tinghög et al., 2013; Bouwmeester et al., 2017). Here, we explore whether the presence and direction of the treatment effect survives variations in the time pressure protocol. The second dimension concerns whether the treatment effect survives when the decisions are taken on a repeat basis rather than in a one-shot game, thus providing variation in the cognitive demand that the decision places on the decider (Duffy and Smith, 2014; Martinsson et al., 2014). This prepares the way to address the second, and key, objective of our paper, namely to investigate whether one or both of the channels through which confusion among subjects could theoretically confound the treatment effect of time pressure on behavior are operational. The nature of the confound matters: Both a heterogeneous treatment effect on confused and unconfused subjects and a link between time-pressure and the likelihood of confusion would weaken the role of cognitive style in explaining the link between time pressure and contributions observed in the existing literature. But while the former would only complicate the identification of the desired treatment effect, the latter would require an entirely new methodological approach.

The empirical strategy of the paper closely follows these two objectives: To meet the first objective, we closely follow the existing literature (Rand et al., 2012, 2014) and randomly assign subjects in a linear PGG either to a baseline condition (BL) in which subjects decide about individual contribution without a time constraint or to a time pressure (TP) condition in which subjects are put under pressure to decide within a tight time limit of seven seconds. The subsequent between-subjects comparison then identifies the causal impact of time pressure on behavior. The TP condition used here employs a standard time pressure protocol that hews closely to the existing literature: Time pressure is only enforced through a modest fixed reduction in the pay-off for violating the time-limit. We implement a first robustness check through an alternative protocol that provides extra incentives (PEI) for deciding within the time limit by drastically increasing the penalty for non-compliance. The second check comes from another protocol that involves real consequences of non-compliance for the game form (PRC): subjects non-compliant with the time limit are shut-out of the PGG and thereby excluded from any payouts. Robustness of the treatment effect with respect to repeat play is implemented through a nine-round repetition that is run after the one-shot game (Duffy and Smith, 2014; Martinsson et al., 2014).

Testing for the potential presence and nature of a confusion confound, our second objective, requires a treatment design that can overcome the problem that disentangling confusion and cooperation in a standard PGG is not possible since positive contributions could be driven by either or both (Andreoni, 1995; Houser and Kurzban, 2002). The design solution proposed in this paper is to combine, in one experiment, the TP treatment with a treatment that has been used previously in order to identify confusion in subjects through a behavioral measure (Houser and Kurzban, 2002; Ferraro and Vossler, 2010; Burton-Chellew et al., 2016). This treatment retains all features of the linear PGG used in the BL and TP condition, except for one difference: Instead of interacting with human partners, subjects interact with computer agents that mechanically contribute a predetermined amount to the public account. The treatment is called *computer condition (CC)* to contrast it with the conventional *human condition (HC)* in which subjects interact with humans. Researchers have previously employed the CC treatment to detect confusion on the basis that its design, if understood by the subject, cannot activate social preferences. When the other party in a PGG is a line of software code drawing on a sequence of pre-determined contributions, contributing to the "public account" does not generate any benefits for the human subject or for any other participant (Houser and Kurzban, 2002; Ferraro and Vossler, 2010).[5] Each subject, whether randomly assigned to the TP or the BL treatment, completes both the HC and the CC, controlling for order effects. This combination of between-subjects and within-subject treatments allows to identify the treatment effect of time pressure on behavior as well as the presence and nature of a confusion confound.

To preview the results, our reinvestigation of the link between cognitive style and contributions

[5] Altruism towards the experiment could be one rival explanation for why subjects contribute in the computer condition. However, Ferraro and Vossler (2010) test whether subjects are motivated to transfer some of their endowment back to the experimenter and dismiss this rival explanation based on two observations: First providing a pay-off table reduced the frequency of positive (confused) contributions in their computer condition from 60 percent of subjects in the baseline to 15 percent in the condition with a pay-off table. Such variation should not occur if subjects were largely unconfused and were simply using the computer condition in order to return some money to the experimenter out of generosity. Second, in ex-post focus group interviews none of their subjects mentioned 'giving back to the experimenter' as a prime motive for contributing in the computer condition. Instead, most answers referred to 'problems to locate the optimal strategy' or 'splitting the endowment' because they were uncertain what would be optimal. This is well in line with the different answers we observe in our ex-post survey in which we asked subjects to state their reasons for contributing in the computer condition. Apart from altruism towards the experimenter Ferraro and Vossler (2010) also discuss and discard other rival explanations that cannot be ruled out by design alone such as the idea that subjects do not understand they are matched with a computer or the idea that subjects understand they are matched with a computer but still behave as if they were matched with a human.

speaks against the conclusion that intuition favors higher contributions (Rand et al., 2012; Cone and Rand, 2014; Rand et al., 2015). The intention-to-treat effect of exposing subjects in the linear PGG with human partners (HC) to time pressure uncovers no evidence of a positive effect on contributions. Rather, subjects are significantly more likely to contribute zero and make weakly lower contributions on average (result 1). Even a less conservative analysis that compares slow (decision time > 7 seconds) and fast (decision time $\leqq 7$ seconds) decisions finds that faster subjects contribute significantly less than slower subjects, even after statistically controlling for potential selection effects (result 2). Taken together, results 1 and 2 lend support to studies such as Tinghög et al. (2013), Verkoeijen and Bouwmeester (2014), Myrseth and Wollbrant (2017), Kocher et al. (2017) and Bouwmeester et al. (2017) that challenge the social heuristics hypothesis, which predicts a positive link between contributions and intuition.

Of the two channels through which cognitive failure could confound time pressure studies of cognitive style, we find only one, the heterogeneous treatment effect, to affect behavior in our experiment. The other channel, by which time pressure could have the undesired side-effect of increasing cognitive failure (Recalde et al., 2014), gives no reason for concern: The likelihood of cognitive failure, measured by positive contributions in the CC, was not higher among subjects exposed to time pressure (result 3). The heterogeneous treatment effect, derives its substance from the high share of confused subjects. In the CC, nearly 50 percent of subjects contribute a positive amount to the public account. Exploiting the within-subjects dimension of our design to identify confused subjects, we find that time pressure in the HC selectively affects only those subjects whose contributions can confidently be interpreted as cooperative: The reduction of average contributions under TP, weakly significant in the full sample, becomes highly significant when we restrict the sample to control for cognitive failure and compliance with the time limit (result 4). This heterogeneous treatment effect could be a key explanation why our reinvestigation of the TP treatment effect, along with those of others, fails to replicate and possibly contradicts earlier findings of the role of cognitive style on cooperation. These main findings on the effects of time pressure are largely robust to two different variations. Variations in the time pressure protocol – either through increased incentives (PEI) or a change in the game form (PRC) – do not affect results (1) and (2) of a link between time pressure and cooperation. However, evidence from the computer condition now suggests that more stringent methods of inducing time pressure could affect subjects' confusion status more heavily (qualifying result 3 to some degree). No heterogeneous treatment effect (result 4) is found in the more drastic PRC protocol, but additional research would be needed to clarify which of the differences in protocol design explains this. Varying the interaction from one-shot to a repeated setting, in which subjects can gain more experience with the task format, also leaves core results (1) and (3) intact. At the same time, the data from the repeat interaction also hint at a new angle, namely at an interaction between time pressure, confusion status, and strategic interaction that must be left to further research.

In the next section we describe the experimental design in more detail. Section 3 spells out our main results. We conclude in section 4 with a discussion and implications.

2 Experimental Design

2.1 Basic Setup

The experimental design combines treatments to identify confused decisions with treatments that speed up decision-making. To disentangle confusion from social preferences we compare two different public good conditions, closely following the design of Houser and Kurzban (2002). In the *human condition (HC)* participants were randomly and anonymously matched into groups of four to participate in a standard PGG. Each participant could decide how to divide an initial endowment (v) of 20 tokens between a private and a public account. A token was worth €0.20 in the private account and contributions (x) to the public account lead to a payoff of €0.10 for all subjects in the group. In other words, each token contributed to the public account was doubled in value (€0.40) and was then split evenly among four group member so that the marginal per capita return (MPCR) equals 0.5. Hence, free-riding is a dominant strategy while full contributions maximize the payoff of the group as a whole. Equation (1) summarizes the linear payoff function for subject i.

$$\pi_i = 0.2(v - x_i) + 0.1(\sum_j^3 x_j + x_i) \qquad \forall i \neq j \tag{1}$$

Positive deviations from this dominant strategy have been attributed to social preferences (e.g., Fischbacher and Gächter, 2010), but also to confusion (Andreoni, 1995; Houser and Kurzban, 2002; Ferraro and Vossler, 2010).

The *computer condition (CC)*, retains all features of the HC with the only difference that the gains from cooperating are removed: Subjects shared their public account with a computer program instead of human interaction partners. Retaining the same basic payoff structure, subjects lost €0.10 for each token contributed without generating additional gains for a group of other participants, as the computer agents did not receive any payoff. Thus, contributions in this condition cannot be attributed to cooperative preferences. Employing this behavioral measure of confusion is central to our design, as it enables us to observe the direct effect of time pressure on the level of confusion, at the moment of decision making. In contrast, an ex-post survey measure would not capture the full effect, as subjects by then would have had additional time to understand the incentives. To analyze the effect of constraining deliberation, the HC and CC were each conducted both in a *baseline setting* (BL) with unconstrained decision time and under *time pressure* (TP). Time pressure was randomly assigned between-subjects. In total we compare four different combinations of treatments: HC-BL, HC-TP, CC-BL or CC-TP.

To assess the individual confusion status needed in a test for heterogeneous treatment effects, we add a within-subjects dimension to the design of Houser and Kurzban (2002). Each subject was observed in one of the HC and in the corresponding CC, controlling for order effects. Under *normal order* the first task was in the HC, while under *reverse order* subjects began in the CC. In both order conditions subjects were informed that there would be a second task, but were uninformed about the specifics of this second task. In order to compare our results to other studies in the literature, we are primarily interested in the outcomes of the one-shot PGGs. However, to assess the role of confusion over time we also conduct a repeated public good game in which subjects are given a possibility to gain additional experience. In each treatment condition subjects also interacted in nine rounds of a repeated public good game with feedback. While subjects knew that they would take additional decisions, the specifics of the repeated protocol were only revealed after the one-shot game. Table 1 summarizes the succession of the different tasks and the corresponding sample sizes. Each condition is described in more detail below.

Table 1: Treatment conditions and order

	Normal Order		**Reverse Order**	
	Baseline	*Time Pressure*	*Baseline*	*Time Pressure*
First Task	**(HC-BL)**	**(HC-TP)**	**(CC-BL)**	**(CC-TP)**
	One-Shot	One-Shot	One-Shot	One-Shot
	Repeated	Repeated	Repeated	Repeated
Second Task	**(CC-BL)**	**(CC-TP)**	**(HC-BL)**	**(HC-TP)**
	One-Shot	One-Shot	One-Shot	One-Shot
	Repeated	Repeated	Repeated	Repeated
Sample Size	N=108	N=112	N=64	N=64

2.2 Computer Condition (CC-BL & CC-TP)

Our behavioral measure of confusion replicates all central elements of the Houser and Kurzban (2002) design. We slightly deviate from their design in the following two aspects: We provide no payoff table in the instructions or on the decision screen in order to rule out that differences in information seeking interact with the effects of time pressure. Furthermore, in each round of the repeated CC subjects did not receive feedback about the actions of the three computer agents prior to, but after stating their own decision in this round. We altered this feature to make the CC more comparable to the HC. Subjects in the CC received the same set of instructions explaining the payoff structure of the standard public good game as subjects in the HC. The only difference was that CC subjects were explicitly informed that their group would consist of three computer agents who (naturally, as they are a computer program) would not receive any payoffs generated through contributions to the public account. On each decision screen we reminded participants of this fact.

To exclude other reasons for contributing in the CC, it is essential that subjects understand the difference between human and computerized interaction partners. Particularly, they should not wrongly assume that the computer was programmed to react to their contribution choices. Therefore, we instructed subjects that the computer agents would contribute predetermined amounts. In order to make this information credible, contributions were written on a concealed poster in the room prior to the experiment and were revealed to subjects at the end of their session. This procedure was described in the instructions before subjects could make any decision. A manipulation check based on two questionnaire items confirms that 92 percent of the subjects understood that they had interacted with a computer program and 93 percent believed that they were not able to influence the computer's contribution.

2.3 Time pressure

The baseline version of the time pressure (HC-TP and CC-TP) treatments strives for high comparability to the existing literature. In its one-shot version, subjects had to decide within seven seconds. This is a slightly stricter limit than in Rand et al. (2012, 2014) and Tinghög et al. (2013). The desired effect of a stricter time limit is to move more subjects closer to the decision time threshold of intuition (Myrseth and Wollbrant, 2017). In the later rounds of the repeated tasks (5-9), the limit was tightened to four seconds to account for possible treatment attenuation through subjects' adaptation to the time constraint over the course of the repetitions. The specific time limits were constructed by subtracting one standard deviation from mean decision time in the first two sessions of the baseline condition. In accordance with the existing literature, subjects were informed about the time constraint only after going through all instructions, right before reaching the decision screen. This procedure prevents subjects in the time pressure condition from changing their behavior on the instruction screen in anticipation of the time constraint. On the decision screen a counter displayed the remaining decision time. In line with the original study by Rand et al. (2012) and its follow-up studies, subjects could violate the time constraint. Without further incentives for complying with the time limit, however, such a design can give rise to high rates of non-compliance and, consequently, to statistical problems (Tinghög et al., 2013) that can be potentially mitigated through compliance incentives. We therefore sacrifice some degree of direct comparability to the treatment effects in Rand et al. (2012) in the interest of internal validity and used a modest incentive that imposed on subjects for each violation of the time constraint a loss of €0.20 of their show-up fee. In this, we follow previous studies on cognitive load in economic games (including cooperation games) that have also incentivised compliance (Cappelletti et al., 2011; Carpenter et al., 2013; Duffy and Smith, 2014). Note that the cost of violating the time constraint are relatively small. Compared to the actual incentives in the public good game they correspond to one token in the private account. This renders a direct effect on behavior in the PGG unlikely.

The baseline treatment is only one of several possible approaches for dealing with the possibility that subjects violate the time limit. Two alternative time pressure protocols serve as a robustness check and explore two different forms of added incentives for compliance. The first alternative (N=56) is equivalent to the normal order TP treatment, but increases the cost for violating time limit through an extra incentive (PEI): Subjects lose their full show-up fee for violating the time constraint. Such extra incentives should increase compliance with the time-limit. A possible downside of the PEI protocol is that subjects might compromise on the quality of decision-making in the service of higher decision speed just in order to escape the sanction, introducing an additional factor into the analysis above and beyond cognitive style (Kocher and Sutter, 2006). Similarly, the more substantial loss incurred for violating the time constraint renders a direct impact on the PGG behavior of subjects who realize that they have just violated the time constraint more likely. The second alternative (PRC) also involves real and even stronger consequences for violating the time limit (N=60). Here, subjects who violate the time constraint are completely shut out of the PGG and all associated earnings. In place of the subject's decision, the protocol then chooses a default contribution from the violator's account to the public account that is an unknown and pre-specified number from the contribution space.[6] For a more extensive discussion on the advantages and disadvantages of incentivising compliance in cognitive load tasks see Duffy and Smith (2014).

2.4 One-shot and repeated decisions

The majority of studies analyzing the effects of time pressure in public good games were conducted in a one-shot environment. To allow for a comparison with these studies, the first decision in our experiment is one-shot as well, in the sense that there was no feedback given regarding the choices of other group members. Subjects were also informed at this point that they would receive a new set of rules after their first decision that would only apply to further decisions.

Experience could play an important role in reducing initial confusion. Therefore, we conduct a repeated version of the same PGG subsequent to the one-shot task. While subjects knew that they would take further decisions in the experiment, they only learned about the specifics of the repeated decisions after stating their choice in the one-shot game. Specifically, they were instructed that there would be nine consecutive rounds within a fixed group of four subjects and that they would receive feedback after each round. Between each decision screen there was a feedback screen displaying the total contributions of their group members. To keep BL and TP comparable, in both conditions the feedback screen was only available for ten seconds after which the next decision screen appeared automatically.

2.5 Experimental procedures

The experiment was conducted at the University of Heidelberg "AWI–LAB" between December 2012 and November 2013. We ran twenty-six sessions of the core TP experiment with sixteen or twelve subjects per session for a total of 348 participants. In October and November of 2017 eight additional sessions for the PRC and PEI treatments were run with a total of 118 participants. The participants were recruited from a standard subject pool of undergraduate and graduate students and randomly assigned to the different treatment conditions. The subjects were from mixed disciplines, including economics (34%). There was a nearly balanced ratio of female

[6] While at first sight an intuitively appealing implementation of real-world time pressure situations, a binding limit in a PGG raises its own complications: It alters the game form of the PGG by allowing subjects the additional option of choosing strategic inaction, i.e. letting the clock run out and the protocol decide on the move. This option introduces unobservables into the analysis – raising statistical problems (Trautmann, 2014) – and thereby reduces comparability with standard TP results used in large parts of the existing literature. An important question for the design of a PRC protocol is how to set the default rule that determines how inaction by one (or more) subject(s) affects contributions to the public account and thereby the earning possibilities of the other group members, especially if there is an omission effect in PGG (Gärtner and Sandberg, 2017). To restrict this concern, we leave the default uncertain. The second problems is that non-violators cannot be sure how many other subjects in their group might violate the time constraint and thus not benefit from any contributions to the public good. Thus, the MPCR of the public good is uncertain and contributions might not lead to an efficiency gain, if at least two of the other group members violate the constraint. Finally, interpreting the contribution choices of violators is harder, because they might realize that their choices are non-consequential once they have violated the time constraint.

(55%) to male (45%) participants.[7] Using ORSEE (Greiner, 2004), subjects who had previously taken part in a public goods experiment at the "AWI–LAB" were excluded from recruitment to the experiment. No participant took part in more than one session of the experiment. Upon arrival, participants were seated at their computer terminal, generated a random password to ensure their anonymity and received a set of general instructions that were read aloud by the experimenter. All other instructions were fully computerized. The decision tasks were implemented using z-Tree (Fischbacher, 2007). During the experiment subjects were only allowed to ask questions in private. Participants were not allowed to communicate with one another. After the decision task, subjects had to complete a set of demographic survey questions and two standardized psychological tests to measure their predisposition for cognitive reflection (Frederick, 2005) and their working memory span (Wechsler, 1955). Furthermore, they were asked to answer incentivised comprehension questions in which they had to state their payoff maximizing strategy and a set of control questions.[8] At the end of the experiment, participants were paid their earnings from one randomly drawn round and task in private. All sessions lasted approximately 75 minutes and participants earned an average of €9.60 (Min.:€4.00;Max.:€15.00), including a show-up fee of €3.[9]

3 Results

We first discuss results from those one-shot PGGs, subjects encountered first in each condition (i.e., Normal Order HC and Reverse Order CC). These outcomes are directly comparable to the evidence in Rand et al. (2012, 2014), Tinghög et al. (2013), Verkoeijen and Bouwmeester (2014), and the recent replication project (Bouwmeester et al., 2017). We then proceed with tests for the robustness of these results to variations of the time pressure protocol and the robustness to repeat interaction.

3.1 One-shot decisions

In Figure 1 we display the effect of time pressure on the distribution of contributions for those subjects making their very first choice in the HC (gray bars: *HC-BL* vs. black bars: *HC-TP*).[10] Despite reproducing essential elements of the design of Rand et al. (2012), Figure 1 does not support their conclusion regarding a tendency to cooperate instinctively in one-shot public good games. Instead, we find a higher incidence of free-riding (BL: 12%; TP: 22%) and a slightly reduced fraction of full contributions (BL: 30%; TP: 25%) in the treatment group. Furthermore, the higher fraction of subjects who split their endowment equally between both

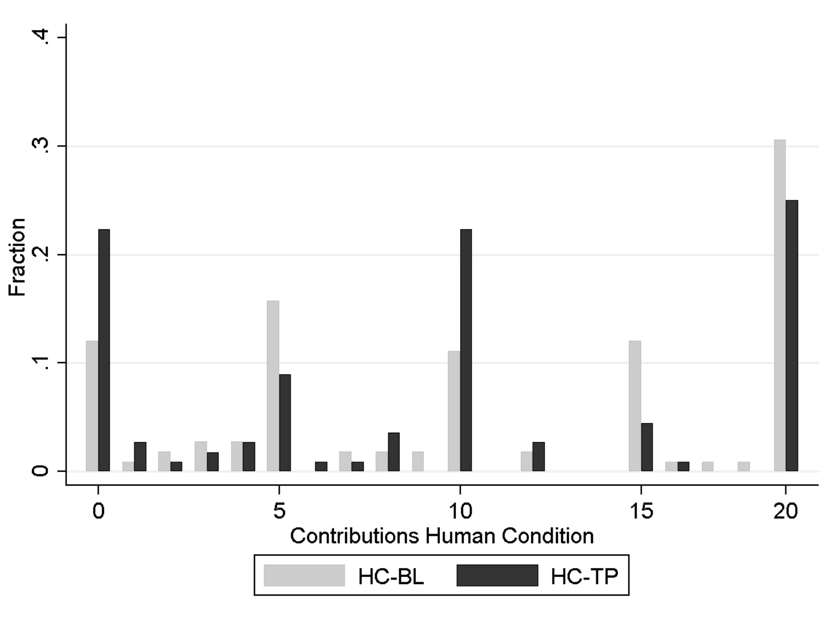
[7]Further summary statistics are contained in Table 15 of the Appendix.

[8]Some of the subjects (N=96) answered these control questions as part of their demographic survey, while others answered them as part of the instructions (N=252). Basic control questions for game form comprehension, could conceivably affect confusion status or behaviour depending on the point during the experiment at which they are administered. As we find no differences in one-shot contributions, both with (M.W. Rank Sum Test: z=0.213, p=0.831) and without applying time pressure (M.W. Rank Sum Test: z=-0.082, p=0.935), we pool observations for the analyses below.

[9]For the additional robustness checks the show-up fee was set to €4.00 to compensate for the higher probability of loosing all earnings due to violating the time constraint.

[10]Remember that these subjects were assigned to the normal order condition, so that behavior cannot be influenced by the subsequent tasks.

Figure 1: Distribution of contributions HC



Notes: This graph shows the distribution of contributions to the public account separately for subjects in the baseline and under time-pressure. Gray bars are used for BL subjects and black bars for TP-subjects. Data from the normal order condition only.

accounts under time pressure (BL: 11%; TP: 22%) could either indicate increased confusion (Ledyard, 1995; Ferraro and Vossler, 2010) or point towards a fairness heuristic (Roch et al., 2000; Cappelen et al., 2016; Capraro et al., 2014). A comparison of mean behavior corroborates these first observations. In the baseline, subjects on average contribute 56 percent of their endowment, which falls into the range typically observed for public good games (Ledyard, 1995). Contributions in the treatment condition are lower, at an average of 47 percent. This difference is weakly significant (M.W. Rank Sum Test: z=1.66, p=0.097) at the ten percent level. We reach an even stronger conclusion, when restricting our analysis to the most extreme forms of defection or cooperation. Time pressure significantly increases free-riding (Chi^2: χ^2=4.07, p=0.044) while it does not affect the fraction of subjects who contribute their full endowment (Chi^2: χ^2=0.85, p=0.357). These results are robust to controlling for additional demographic (age, sex, risk aversion, correct answer to control question) and psychometric (time spent reading the instructions, test scores from cognitive reflection test, and working memory test) variables, as shown by multiple regressions in Table 7 of the Appendix. Taken together, we therefore reject the hypothesis of intuitive cooperation and state the following result:

Table 2: Instrumental variable estimates of the effects of fast decision making in the human condition

	(1) Contributions	(2) Contributions	(3) Free-Riding	(4) Free-Riding
Panel A: Second Stage (DV: Contributions)	**HC**	**HC**	**HC**	**HC**
Response Time (Log10 Sec.)	3.654**	4.164***	-0.441***	-0.489***
	(2.36)	(2.71)	(-2.67)	(-2.81)
Age (Years)		-0.224		-0.009
		(-0.81)		(-0.26)
Sex (1=Male)		3.510**		-0.139
		(2.45)		(-0.85)
Risk Aversion (1-11)		0.214		-0.003
		(0.70)		(-0.09)
Unconfused (1=Yes)		-2.192		0.240
		(-1.48)		(1.45)
Cognitive Reflection Score (0-3)		1.965***		0.045
		(2.90)		(0.60)
Readingtime Instructions (Log10 Sec.)		3.104*		-0.048
		(1.84)		(-0.24)
Working Memory Score (0-1)		2.760		-0.509
		(0.59)		(-1.09)
Constant	2.209	-10.36	0.102	0.632
	(0.63)	(-1.05)	(0.28)	(0.53)
Observations	348	335	348	335
Panel B: First Stage (DV: Response Time)				
Treatment(I): Normal Order + Time Pressure (1=Yes)	-0.855****	-0.814****	-0.855****	-0.814****
	(-12.49)	(-11.78)	(-12.49)	(-11.78)
Treatment(II): Reverse Order (1=Yes)	-0.583****	-0.610****	-0.583****	-0.610****
	(-6.67)	(-7.36)	(-6.67)	(-7.36)
Treatment(III): Reverse Order + Time Pressure (1=Yes)	-1.324****	-1.327****	-1.324****	-1.327****
	(-16.93)	(-15.77)	(-16.93)	(-15.77)
Age (Years)		-0.001		-0.001
		(-0.12)		(-0.12)
Sex (1=Male)		-0.019		-0.019
		(-0.36)		(-0.36)
Risk Aversion (1-11)		0.005		0.004
		(0.37)		(0.37)
Unconfused (1=Yes)		0.068		0.068
		(1.25)		(1.25)
Cognitive Reflection Score (0-3)		-0.029		-0.029
		(-1.18)		(-1.18)
Readingtime Instructions (Log10 Sec.)		0.243****		0.243****
		(3.44)		(3.43)
Working Memory Score (0-1)		-0.327**		-0.327**
		(-2.16)		(-2.16)
Constant	2.830****	2.167****	2.829****	2.167****
	(48.84)	(6.18)	(48.71)	(6.18)
F-Statistic First Stage	108.66	34.84	108.66	34.84

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Notes: Specifications (1) and (2): Tobit-IV maximum likelihood estimation to account for censoring from below (0) and above (20). Specifications (3)-(4): Probit-IV maximum likelihood estimation. t-statistics in parentheses. Robust standard errors. Estimates for the pooled sample. Additional robustness checks in tables 8 and 9 of the online appendix.

Result 1: ***There is no evidence for a greater tendency to contribute to the public account under time pressure. Instead, time pressure significantly increases the incidence of zero contributions and weakly decreases average contributions.***

As expected, time pressure induces subjects in the treatment condition to spend significantly less time on the first decision screen (M.W. Rank Sum Test: z=10.48, p<0.001). Median response times are 16 seconds in the BL and 7 seconds in the TP condition.[11] However, only 57.1 percent of subjects under time pressure make their decision within the set time limit, whereas 7.4 percent of subjects in the baseline decide within seven seconds. Decisions from subjects who chose to spend more time on the decision screen are less informative for identifying the effects of intuition (Myrseth and Wollbrant, 2017). Consequently, the more conservative intention-to-treat effect of forced intuition (i.e., the effect of treatment assignment) we report above corresponds to a weighted average of the zero (or reduced) effect on non-compliers and the true treatment effect on compliers (Bloom, 1984). In other words, it most likely understates the true impact of constraining deliberation. Therefore, we now adjust previous results for compliance. When we simply compare fast subjects (decision times $\leqslant 7$ seconds, henceforth: fast) to slow subjects (decision times > 7 seconds, henceforth: slow) the negative effect of constrained deliberation on the size of contributions as well as on the probability to contribute at all, increases in size and significance. Fast subjects contribute 40 percent of their endowment and slow subjects 57 percent (M.W. Rank Sum Test: z=3.11, p<0.01). Similarly, 31 percent of fast subjects free-ride as compared to 10 percent of slow subjects (Chi^2: χ^2=16.12, p<0.001). Before this result can stand, the analysis needs to account for potential selection effects, since fast and slow subjects might differ in observable or unobservable ways (Tinghög et al., 2013). The data generated by our experiment do not point towards the presence of selection bias on the basis of observed characteristics.[12] The only exception could be working memory capacity which is significantly higher for fast subjects (M.W. Rank Sum Test: z=-4.63, p<0.001), but which is already at increased levels for subjects randomized to the TP condition (M.W. Rank Sum Test: z=-2.88, p<0.01). There is, however, still the possibility that some unobservable subject characteristic is correlated with both response times and contribution behavior. To account for potential problems resulting from self-selection, we follow Imbens and Angrist (1994) and Angrist et al. (1996) and use assignment to one of the treatment conditions as an instrument for potentially endogenous response times. By design, treatment assignment is random and hence exogenous to the degree that treatment assignment was randomized on a session level in a balanced fashion (Fréchette, 2012)[13]; still, treatment assignment is highly correlated with faster decisions. Table 2 displays estimates from four instrumental variable regressions. Tobit regressions (1) and (2) take contributions as the dependent variable, probit regressions (3) and (4) the decision to free-ride or not. First stage regressions (Panel B) show that random assignment to treatment significantly decreases response times in each case relative to the unconstrained baseline. Furthermore, two psychometric variables are correlated with fast decision making. Subjects with a higher working memory capacity and subjects who spend less time reading the instructions make faster choices on the actual decision screen. Plausibly, a better ability to remember the details of the task speed up decision making. Second stage regressions in Panel A show the main effect of interest. Across specifications (1)–(4) faster decisions lead to significantly lower contributions and significantly

[11] These values are computed for subjects in the normal order condition. An overview over the full distribution of response times across all treatment conditions is given in Table 6 of the Appendix.

[12] An overview over subject characteristics by treatment and compliance status is given in Table 15 of the Appendix.

[13] Sessions were run across several times and weeks. Apart from some design-related limitations (i.e. the requirement to run two of the baseline sessions as the first sessions to calculate the time limits for later time pressure sessions) we made sure that treatment randomisation was balanced across all possible session times.

more free-riding. Thus, adjusting for potential selection-bias, IV results confirm a positive effect of more deliberation on contributions: a ten percent increase of time spent on the decision screen increases contributions by 0.35 tokens. This positive link is robust to controlling for additional demographic and psychometric variables in regressions (2) and (4). Male subjects contribute significantly more. Confusion status, assessed by a simple survey question is not correlated with contribution behavior. Yet, two of the included psychometric variables are related to the experimental outcome. The amount of time subjects spend on the instructions screen is included as a general measure of their engagement and reading speed. Furthermore, it could be related to the amount and kind of information subjects acquire when reading the instructions. It has been shown (Fiedler et al., 2013b) that subjects who care more about the payoffs of other participants acquire more information about the payoff structure and consequently might spend more time on reading the instructions. This interpretation would be in line with the weakly positive relationship shown in regression (2). Subjects could also differ in their propensity to rely on their intuition. We control for these differences by scores from a cognitive reflection test (CRT) (Frederick, 2005). In line with the main treatment effect we find that subjects who are more prone to rely on a deliberative cognitive style, as measured by the CRT, contribute more to the public good.[14] Both psychometric measures are not associated with the rate of free-riding (4).

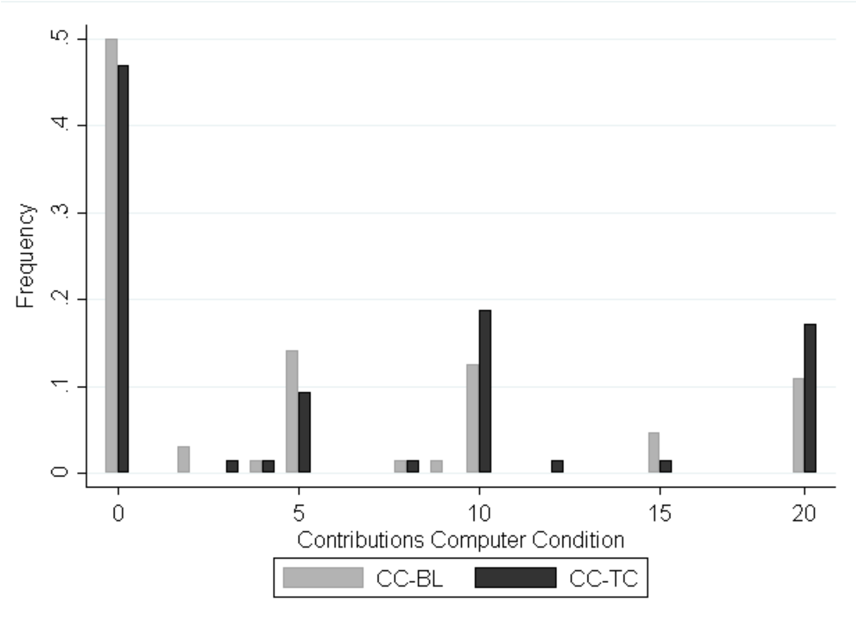
Result 2: ***Faster subjects contribute significantly less to the public account than slower subjects. After controlling for potential selection effects, we still find support for a causal link between more deliberation and higher contributions.***

We continue by analyzing choices of those participants who took their first one-shot decision in the computer condition (CC). Figure 2 compares the distribution of CC contributions from the reverse order condition between subjects in the baseline and subjects in the time pressure condition. Thus, it illustrates how constraining the use of a reflective cognitive style affects behavior in a situation of comparable complexity to the HC, but in which gains from cooperation cannot motivate behavior. Time pressure only slightly increases the occurrence of confused contributions: fewer participants stick to their dominant strategy of contributing zero tokens (BL: 50%; TP: 47%), whereas there are more participants who give up half of (BL: 12%; TP: 19%) or even their full (BL: 11%; TP: 17%) endowment. None of these differences reaches statistic significance at conventional levels. This continues to hold when we adjust results for compliance with the time constraint. Fast subjects are neither significantly less likely to contribute zero (Chi^2: χ^2=0.20 $p = 0.648$), nor do they contribute more on average (M.W. Rank Sum Test: z=-0.55, $p = 0.791$).[15] Therefore, in contrast to the concerns raised by correlational evidence in Racalde et al. (2014), we conclude that taxing participants' deliberative capacities by applying time pressure does not increase confusion levels in our setting.

[14] In a companion paper (Lohse, 2016) we explore and interpret this relationship more thoroughly using parts of the same dataset.

[15] Results from the corresponding IV regressions confirm this finding and are available on request.

Figure 2: Contribution Frequencies CC



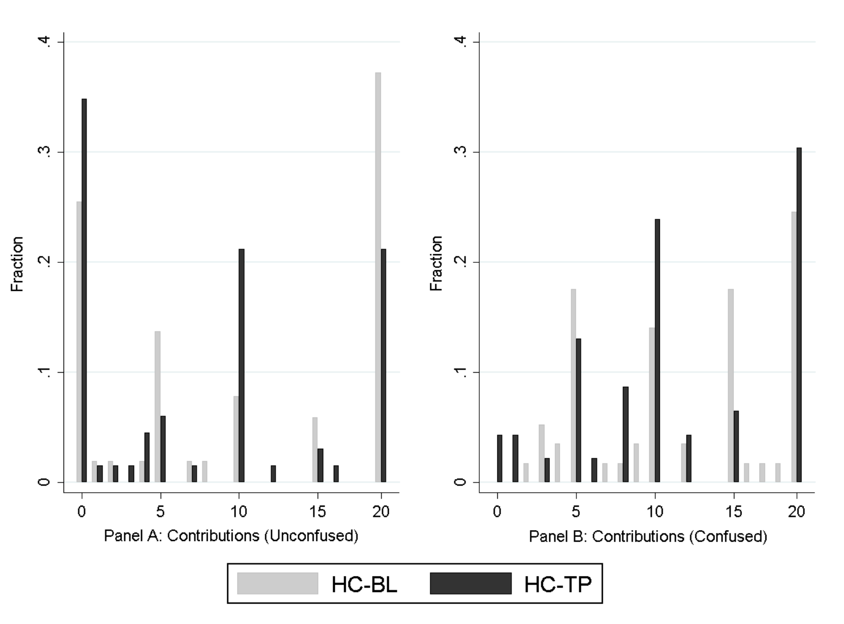
Notes: This graph shows the distribution of contributions to the public account in the computer condition (RO) separately for subjects in the baseline and under time-pressure. Gray bars are used for BL subjects and black bars for TP-subjects.

Result 3: ***In the one-shot CC, we observe no effect of time-pressure on contributions.***

As in Houser and Kurzban (2002) and Ferraro and Vossler (2010), approximately half of the participants in the CC contribute positive amounts, despite the fact that this reduces their own payoffs without benefiting any other group member. This substantial presence of confusion could complicate the interpretation of the link between contributions and cooperation. Only for subjects who show no sign of confusion in the computer condition a treatment effect in the HC can confidently be attributed to a change in cooperative behavior. Furthermore, time pressure could affect subjects selectively by confusion status. We exploit the within-subjects structure of our data to devise two different tests for these potential concerns.

In our test for heterogeneous treatment effects we split the sample into "confused" and "unconfused" subjects. We do so by sorting a participant into the "confused" bin if we observe positive contributions in the one-shot game of the CC and into the "unconfused" bin if we observe zero contributions. Note that this classification is conservative in the following sense: At the point where subjects make their first contribution choice in the computer condition (normal order) they will already have made ten choices in the preceding human condition. Hence, they have already gained some experience with the task that might have resolved some of their initial confusion. Thus, our classification constitutes a lower bound of the actual confusion present during the first decision in the HC. Panels A and B of Figure 3 compare the effect of time

Figure 3: Distribution of contributions by confusion status



Notes: This graph shows the distribution of contributions to the public account separately for subjects in the baseline and under time-pressure using observations from the normal order condition. Gray bars are used for BL subjects and black bars for TP subjects.

pressure on contributions between subjects in the "confused" and "unconfused" bins.[16] Two observations stand out clearly: First, for baseline subjects the distribution of contributions differs by their confusion status. None of the confused subjects contribute zero tokens. Confused subjects are not more cooperative in general, as they are also less prone to contribute their full endowment. Instead, they more frequently choose a contribution from within the contribution range.[17] Overall, the average contributions of confused subjects are significantly higher than for unconfused ones (M.W. Rank Sum Test: z=-3.06, p<0.01). Second, the effect of time pressure appears to work in opposite directions, by confusion status. For unconfused subjects time pressure increases free-riding and decreases full contributions. For confused subjects time pressure slightly increases full contributions.

When testing for a heterogeneous treatment effect, we find time pressure to reduce average contributions only among unconfused subjects. This effect gets stronger when we adjust results for compliance to time pressure. On average, fast subjects contribute significantly less (fast: 34.6%; slow: 53.8%) if unconfused (M.W. Rank Sum Test: z=2.99, p<0.01), but approximately the same average amount if confused (M.W. Rank Sum Test: z=0.13, p=0.89). Instrumental variable regressions in Table 3 confirm that these results are not driven by selection effects.

[16]We display results for the normal order condition to allow for a clean comparison to Figure 1. Results, however, do not differ, when using pooled data.

[17]Remember, intermediary contributions are not consistent with the predictions of many standard social preference models, which posit that decision makers either contribute nothing or their full endowment, depending on the strength of their other-regarding concerns. Therefore, it would not be surprising if intermediary contributions were more common among confused participants, who might mistakenly think that contributing half of the endowment equalizes payoffs.

Table 3: Instrumental variable estimates of the effects of fast decision making separated by confusion status

	Unconfused		Confused	
	(1)	(2)	(3)	(4)
	Contributions	Contributions	Contributions	Contributions
Second Stage (DV: Contributions)				
Response Time (Log10 Sec.)	7.864**		2.495*	
	(2.23)		(1.74)	
Response Time ≦ 7 Sec. (1=Yes)		-12.190**		-4.030
		(-2.29)		(-1.56)
Age (Years)	-0.455	-0.278	-0.053	-0.043
	(-0.74)	(-0.45)	(-0.21)	(-0.17)
Sex (1=Male)	4.789	5.170*	3.421**	3.401**
	(1.53)	(1.65)	(2.54)	(2.50)
Risk Aversion (1-11)	-0.240	-0.311	0.464*	0.506*
	(-0.35)	(-0.47)	(1.76)	(1.86)
Cognitive Reflection Score (0-3)	2.945*	2.578*	1.956***	1.882***
	(1.80)	(1.67)	(3.20)	(3.04)
Readingtime Instructions (Log10 Sec.)	4.846	5.761	1.210	1.645
	(1.31)	(1.60)	(0.82)	(1.10)
Working Memory Score (0-1)	8.443	11.45	1.022	2.359
	(0.80)	(1.09)	(0.25)	(0.50)
Constant	-26.10	-10.96	-3.942	0.982
	(-1.21)	(-0.54)	(-0.45)	(0.12)
Observations	170	170	165	165
First Stage F-Statistic	19.40	12.12	19.03	14.33

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Notes: Tobit-IV maximum likelihood estimation to account for censoring from below (0) and above (20). Specifications (1)-(2): subjects classified as unconfused. Specifications (3)-(4): subjects classified as confused. t-statistics in parentheses. Robust standard errors. Estimates for the pooled sample. First stage available on request.

Specifications (1) and (2) contain estimates for unconfused subjects. Potentially endogenous response times (1) or a dummy indicating fast decisions below seven seconds (2) are again instrumented by exogenous treatment assignment. In both specifications faster decisions lead to significantly lower contributions. In contrast, specifications (3) and (4) show that subjects classified as confused are largely unaffected by their decision speed. Those who decide within seven seconds (4) do not differ from slower decision makers in their contribution behavior. The effect for a continuous response time variable (3) remains weakly significant, but is quantitatively much smaller that the corresponding effect (1) for unconfused subjects. These findings are robust to switching to a more regressive criterion by which we sort participants into the "confused" bin. When only sorting subjects into the "confused" bin, because they were unable to identify the strategy that would have maximized their own payoff in a control question and additionally made a positive contribution in the CC, we again find that time pressure selectively affects unconfused subjects. From these observations we state the following result:

Result 4: ***Those subjects who both understand the incentive structure and decide fast under time pressure can be said to cooperate less. Those giving reason to doubt whether they understand the incentive structure of the PGG are largely unaffected by time pressure***

Overall, our results from the one-shot games show that constraining deliberation by applying time pressure reduces contributions to the public account. Contrary to our initial expectations, we do not find time pressure to directly increase confusion in the CC. However, there is evidence that time pressure selectively affects participants who display no signs of confusion in the CC. This points towards one potential explanation why we, in line with Tinghög et al. (2013) and Verkoeijen and Bouwmeester (2014), fail to replicate evidence on an intuitive predisposition towards cooperation (Rand et al., 2012, 2014). In other words, for those subjects for which contributions can be safely equated with cooperation, because they do not show signs of confusion in the CC, our results suggest that reflection and not intuition is driving cooperative behavior.

3.2 Robustness to Variations in the Time Pressure Protocol

The time pressure protocol under which results (1) through (4) are derived is known to generate incomplete compliance with the time limit across subjects, justifying the study of alternative protocols such as PEI and PRC. These alternatives provide a test of robustness as well as new methodological insights into the use of different time pressure protocols, even though their results are by design less comparable to those of the original papers of Rand et al. (2012, 2014), Tinghög et al. (2013), Verkoeijen and Bouwmeester (2014) and Bouwmeester et al. (2017).

Comparing decision times under PEI and PRC with TP, we do not observe large differences between the different time pressure protocols. Median decision time under TP was 7 seconds. In the PEI treatment, i.e. with extra incentives for compliance, we observe a slightly lower median response time of 6 seconds. In the PRC treatment, in which subjects violating the time limit are excluded from PGG benefits, we observe a slightly increased median response time of 8 seconds. In each of the time pressure treatments, decision times are significantly smaller than in the baseline, in which the median subject required 16 seconds to state their preferred contribution level. Compared to the standard time pressure treatment a significantly higher fraction of subjects (71.4%) complies with the time restriction in the PEI treatment (Chi^2 Test: p=0.072). Despite higher incentives, there is is a significantly lower fraction (40%) of compliance in the PRC treatment (Chi^2 Test: p=0.032)[18].

Table 7 in the appendix contains estimates for the average treatment effects of both additional time pressure conditions on contributions. In line with result 1, and contradicting earlier research on the social heuristics hypothesis (Rand et al., 2014), we do not find that subjects in the PEI or PRC condition contribute significantly higher amounts than subjects in the baseline condition. Compared to the baseline the average rate of contributions is slightly lower (54.7%) in the

[18]If we also include subjects who have missed the time limit by just one second in the group of compliers, compliance rates in all time pressure treatments are relatively high with 78 percent in the TP, 83 percent in the PEI and 64 percent in the PRC treatments. Especially in the PEI treatment the rate of compliance is significantly higher than in most of the existing literature, where compliance rates in most studies were in the range of 34 percent (Bouwmeester et al., 2017) to 52 percent (Rand et al., 2012).

PEI treatment and slightly higher in the PRC (59.8%) condition. The rate of free-riding also increases compared to the baseline in both the PEI (16%) and PRC (15%) treatment, although not as significantly as in the standard time pressure condition.

We next include subjects from the PEI and PRC conditions into a pooled comparison of slow and fast subjects. We continue to find that fast subjects contribute significantly (M.W. Rank Sum Test: p=0.0594) less to the public account than slow subjects, who require more than 7 seconds for their decision (Fast: 48.75%, Slow: 56.7 %). A separate comparison of each of the additional treatments provides a more nuanced picture. In the PRC condition, fast subjects contribute slightly less than slow subjects (M.W. Rank Sum Test: p= 0.456). In the PEI condition, fast subjects contribute more than slow subjects, but not significantly so (M.W. Rank Sum Test: p= 0.1882). One possible interpretation of this deviation from the TP condition is that incentives for compliance might affect choices in the game: Subjects in PEI who realise that they have violated the time limit and thus lost their show-up fee might try to recover some of their losses by lowering their contribution upon additional deliberation. This is one of the drawbacks of using significant incentives for compliance, as they have a higher likelihood of affecting behaviour in the game. Contributions in the PRC treatment that are made after the time-limit are hard to interpret because subjects might realise that their choices are non-consequential; discarding these observations, however, would lead to statistical selection effects (Trautmann, 2014). In table 10 of the Appendix we provide IV estimates of decision time on contribution behaviour similar to those that inform result 2. Including the additional observations from the PEI and PRC conditions into this analysis does not change the previous conclusions summarized in result 2. After controlling for selection effects, we still find support for a causal link between more deliberation and higher contributions as well as lower rates of free-riding.

Enforcing time pressure more stringently might also have the undesired effect of affecting the level of confusion. Our additional time pressure conditions can only shed partial light on this question, as they have only been conducted as the second task in the normal order condition while our previous discussion of the computer condition has primarily relied on observations from the reverse order condition, in which the CC was the first task. In the PEI treatment, subjects contribute 35 percent of their endowment and in the PRC treatment 36.75, as compared to 24.5 percent in the baseline. These differences, indicating an increase in confusion levels, are (weakly) significant (M.W. Rank Sum Test, PEI: p=0.0832, PRC: p=0.0531). The rates of confusion measured by the probability of a positive contribution are at 61 percent for PEI and 65 percent for PRC. These rates are higher than in the baseline (53 percent), although not significantly so (Chi^2, PEI: p=0.332, PRC: p=0.125). Jointly, these observations only lend partial support to result 3. On a methodological level they indicate that strongly enforcing a time limit might have the undesired effect of inviting additional confusion, which in turn could endanger the clear identification of changes in cooperative behaviour based on observed contributions.

In Result 4 we describe a heterogeneous treatment effect suggesting that time-pressure selectively affects unconfused subjects. We observe a similar effect in the PEI treatment. On average, subjects in the PEI treatment contribute less (42%) than subjects in the baseline (51.125%) if unconfused, but slightly more if confused (62.9% compared to 60%). When simply comparing fast and slow subjects these differences become larger. Fast subjects in the PEI contribute less if unconfused (40 %), but significantly (M.W. Rank Sum Test, z = -2.474, p=0.0134) more if confused (75 %). For the PRC condition we do not observe a similar pattern, with confused (fast: 57 %, slow: 62%) and unconfused (fast: 58 %, slow: 61%) subjects contributing similar amounts independent of their decision speed. Regression results in Table 11 of the appendix demonstrate that the heterogeneous treatment effect persists when including observations from the PEI and PRC treatments into a pooled comparison: longer deliberation times lead to higher contributions among unconfused subjects, while not significantly affecting the contributions of confused subjects. Jointly these findings lend further support to the observation that confusion is an important moderator in studies that are interested in the effects of cognitive style on cooperation.

3.3 Robustness to Repetition

Subsequent to each one-shot decision, participants remained in their assigned treatment conditions (HC-BL, HC-TP, CC-BL, CC-TP) and took decisions in nine rounds of a finitely repeated PGG. Therefore, in total, every participant completed two distinct one-shot PGG and two distinct repeated PGG, one of each in the human and one in the corresponding computer condition (compare Table 1). Prior to taking their first decision of the repeat PGG, participants were instructed that they would receive feedback regarding the total contributions of the other three group members (predetermined total contributions of the three computer agents) at the end of each round. These additional observations from the repeated games allow us to explore two issues which have not been addressed in the previous literature on time pressure in the PGG: First, by comparing aggregate behavior across the different conditions, we assess the persistence of treatment effects to repetition. Second, analyzing the evolution of individual decisions across different rounds, we evaluate how confusion, experience, and time pressure interact to shape strategic behavior.

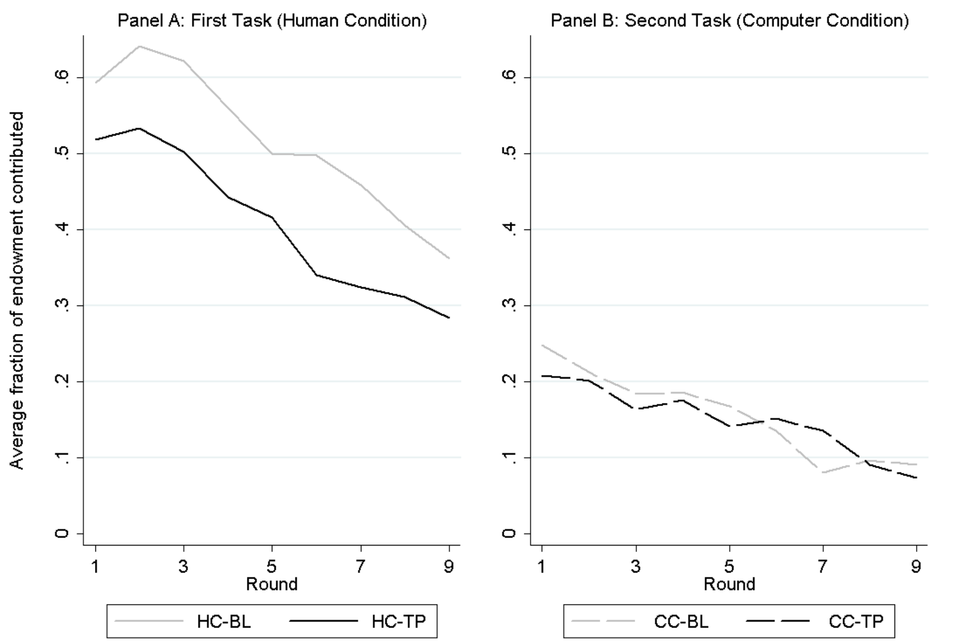
Table 4: Contributions averaged over nine rounds across treatment conditions

	(I) Normal Order		**(II) Reverse Order**	
	Baseline	*Time Pressure*	*Baseline*	*Time Pressure*
First Repeated Public Good Game	HC-BL (N=108)	HC-TP (N=112)	CC-BL (N=64)	CC-TP (N=64)
Contribution Average (% of endowment) (s.d.)	0.52 (0.28)	0.41 (0.27)	0.18 (0.21)	0.20 (0.22)
Second Repeated Public Good Game	CC-BL (N=108)	CC-TP (N=112)	HC-BL (N=64)	HC-TP(N=64)
Contribution Average (%of endowment) (s.d.)	0.16 (0.21)	0.15 (0.21)	0.35 (0.31)	0.42 (0.31)

On the first point, table 4 presents summary statistics on contribution rates averaged over all rounds. Summarizing the more detailed between-subjects comparisons that follow below, the treatment effects detected in the one-shot games appear robust to repetition. Proceeding from the top, the upper row summarizes decisions from those repeated games that subjects encountered first under each condition. Subjects in these games have only been exposed to a limited amount of experience, namely a single decision in the preceding one-shot game. As in result 1, we find a significantly negative effect of time pressure in the HC (Group-level M.W.

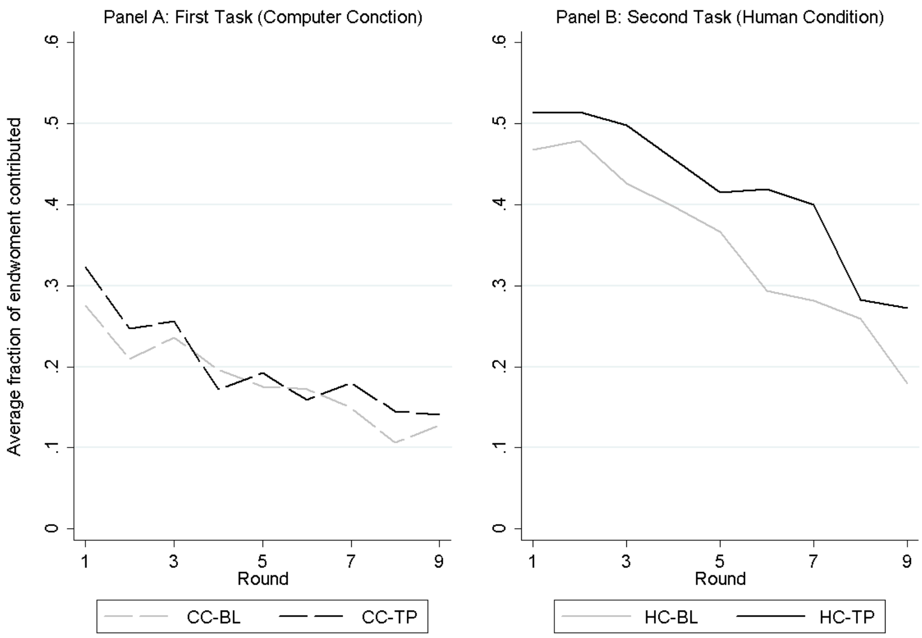
Rank Sum Test: z=2.20, $p = 0.028$) and, as in result 3, no significant effect in the CC (Group-level M.W. Rank Sum Test: z=-0.62, $p = 0.534$). Moving to the second repeated games (i.e., more experienced subjects) displayed in the bottom row, time pressure affects neither average contributions in the HC (Group-level M.W. Rank Sum Test: z= -0.68, $p = 0.498$), nor in the CC (Group-level M.W. Rank Sum Test: z= 0.24, $p = 0.814$). This finding is consistent both with a general experience effect, which makes subjects become more accustomed to deciding under a time limit, and a specific experience effect of the CC. Comparing average contributions in the HC across the two different order conditions indicates that contributions in the reverse order condition are significantly lower without time pressure(Group-level M.W. Rank Sum Test: z= 2.412, $p = 0.0159$), but do not differ significantly with time pressure (Group-level M.W. Rank Sum Test: z= 0.183, $p = 0.8548$). This could point to an effect of time pressure on learning: Subjects in the HC-BL condition may be better able to learn from their experiences in the preceding CC than subjects under time pressure (HC-TP). Further experimental work will be required to test the competing hypotheses that could explain these patterns.

Figure 4: Round-wise average contributions (Normal Order)



Moving towards the evolution of behavior across the nine rounds, figure 4 displays how average contributions (as a fraction of endowment) develop over time in each of the four conditions conducted under normal task order. Panel A shows contributions in the human condition (First Task: HC-BL and HC-TP), Panel B contributions in the subsequent computer condition (Second Task: CC-BL and CC-TP). Across all four conditions, contributions exhibit the typical convergence towards the equilibrium. In the HC, the share of zero contributions nearly doubles from 23 percent in the first round to 41 percent in the final round. Comparing the game dynamics between TP and BL, we again find no evidence for intuitive cooperation, in line with result 1:

Figure 5: Round-wise average contributions (Reverse Order)



Participants under time pressure contribute less in each of the nine rounds and converge towards equilibrium at a comparable speed. Irrespective of time pressure, we observe no pronounced end-game effects in the last round. Moving to the CC treatment, the game dynamics there are also in line with result 3 in that time pressure does not appear to affect the level of confusion. In the first round, 46 percent of subject contribute a positive amount compared to 25 percent in the last round. The decline of contributions is steeper in the HC than in the CC. This lends support to the interpretation that declining contributions in the CC mostly represent a reduction of confusion, while declining contributions in the HC could additionally be due to "frustrated attempts" (Andreoni, 1995, p.892) at unreciprocated cooperation.

Figure 5 displays the contribution patterns for the four conditions conducted in reverse task order. In Panel A, we show average contributions from the computer condition (First Task: CC-BL and CC-TP) and in Panel B contributions from the subsequent human condition (Second Task: HC-BL and HC-TP). For subjects taking their first repeated decisions in the CC, result 3 again appears to hold: There is no evidence for increased confusion under time pressure or a slower convergence towards zero contributions. However, consistent with learning, the initial level of confused contributions is slightly higher and the subsequent decline steeper than in the corresponding rounds from the CC conducted under normal order. Similarly, subjects deciding in the HC-BL after completing the CC start at a lower level of contributions when there is no time limit. A significant restart effect between the CC and the HC suggests that learning accounts only partially for the decline of contributions (Andreoni, 1988). Finally, and in line with result 1, in the HC there is no significant difference in average contributions between subjects in the baseline and subjects under time pressure. Thus, subjects who are more familiar with the task and the time pressure manipulation display neither an intuitive tendency to cooperate or to defect.

Random effects regressions shown in Table 12 of the online appendix confirm the observations from Figures 4 and 5. Summarizing their main results, which use the choice data pooled across both task orders, we find, first, that time pressure significantly reduces contributions in the normal order condition by approximately two tokens. Second, in the reverse order treatment subjects significantly reduce their contributions by three tokens compared to the normal order treatment. Third, in the reverse order treatment subjects under time pressure contribute roughly 0.7 tokens more than subjects deciding without time pressure, a difference that is not significant at conventional levels (Wald test, p=0.638). Together, these three observations lead to

Result 5: ***In line with Result 1, time pressure does not increase contributions in the repeated games. Instead, it significantly reduces contributions in the normal order condition. Time pressure also does not significantly increase contributions in the reverse order condition.***

The same regressions suggest that time pressure has little effect on contributions in the computer condition, in line with result 3: There is no significant effect in the normal order condition. In the reverse order condition, subjects under time pressure contribute approximately one token more than subjects deciding without time pressure, a difference that is not significant (Wald test, p=0.3625). We summarize in

Result 6: ***In line with Result 3, there is only weak evidence that time pressure increases confusion in the CC of the repeated games. Furthermore, time pressure marginally affects the rate at which confusion is reduced.***

Testing the treatment effects of time pressure on contributions for robustness to repeat play is not the only way to use the data from the repeated game. They can also shed light on possible reasons for this robustness by examining how deliberation and confusion interact to shape strategic behavior in a repeated PGG setting. Evidence from the prisoner's dilemma (Milinski and Wedekind, 1998; Duffy and Smith, 2014) demonstrates that constraining deliberation via cognitive load can have an impact on strategic behavior. Specifically, these studies find that subjects under cognitive load are less able to condition their own decisions on their partner's past decisions. Testing whether time pressure similarly reduces strategic behavior in a PGG requires a method for identifying strategic behavior. The obvious candidate for the latter is conditional cooperation, the most frequent type of strategic behavior in a typical PGG (Fischbacher et al., 2001; Chaudhuri, 2011). Evidence from one-shot games such as in Rand et al. (2012, 2014) and similar papers cannot be used to disentangle whether time pressure affects conditional or unconditional cooperation in their setting (Gächter, 2012).[19] Data from repeated PGG, on the other hand, can provide the required evidence.

[19]Based on the strategy method, Nielsen et al. (2014) provide correlational evidence that conditional cooperation is faster than defection. However, as for other correlational studies based on endogenous response times, this relationship cannot be interpreted as causal evidence that intuition favors conditional cooperation.

To test this hypothesis and identify conditional cooperation in our data, we follow the empirical strategies described in Croson et al. (2005), Croson (2007), and Ashley et al. (2010). We estimate a set of panel regressions in which we model how individual contributions change from round t-1 to round t. This first difference in contributions can depend on the behavior of the other group members in round t-1. In theory, a conditional cooperator will contribute more in round t, if his contributions are below the group average in round t-1. Similarly, he will reduce his contributions in round t, if his contributions exceed the group average in t-1. We define two dummy variables to capture this relationship in our regression framework. Subjects contributing the same amount as the group average serve as the reference category. A robustness check reported in Table 14 of the online appendix defines these groups based on tertiles.

Table 5: Repeated decisions: conditional cooperation

	HC: Full Sample		HC: Unconfused		HC: Confused	
	(1)	(2)	(3)	(4)	(5)	(6)
	Δ Contributions	Δ Contributions	Δ Contributions	Δ Contributions	Δ Contributions	Δ Contributions
Treatment(I): Normal Order + Time Pressure (1=Yes)	-0.093	1.444	-0.127	2.843*	0.066	-2.046
	(-0.50)	(1.09)	(-0.54)	(1.70)	(0.17)	(-1.03)
Treatment(II): Reverse Order (1=Yes)	-0.215	2.091*	-0.084	3.466***	-0.543	-1.626
	(-1.03)	(1.84)	(-0.32)	(2.75)	(-1.27)	(-0.85)
Treatment(III): Reverse Order + Time Pressure (1=Yes)	-0.202	0.389	-0.175	1.720	-0.341	-2.679
	(-0.87)	(0.36)	(-0.56)	(1.50)	(-0.79)	(-1.38)
Above Group Average in t-1 (1=Yes)	-2.912****	-1.441	-2.716****	-0.0119	-3.330****	-5.140***
	(-7.20)	(-1.33)	(-4.63)	(-0.01)	(-6.98)	(-2.91)
Below Group Average in t-1 (1=Yes)	1.654****	3.302***	1.569***	3.887****	1.861****	1.447
	(4.34)	(3.19)	(2.90)	(4.35)	(4.10)	(0.88)
Round (1-9)	-0.144***	-0.202**	-0.147***	-0.217**	-0.131**	-0.195
	(-3.28)	(-2.10)	(-2.74)	(-1.99)	(-2.14)	(-1.53)
Treatment(I)*Above		-2.229*		-4.229**		2.109
		(-1.70)		(-2.57)		(1.09)
Treatment(II)*Above		-2.087*		-4.001***		2.395
		(-1.70)		(-3.28)		(1.21)
Treatment(III)*Above		-0.832		-2.463*		2.738
		(-0.67)		(-1.76)		(1.47)
Treatment(I)*Below		-2.099		-3.547**		1.379
		(-1.58)		(-2.17)		(0.75)
Treatment(II)*Below		-2.502**		-3.327***		-0.168
		(-2.23)		(-3.23)		(-0.09)
Treatment(III)*Below		-1.700		-2.226*		-0.132
		(-1.46)		(-1.71)		(-0.07)
Treatment(I)*Round		0.101		0.142		0.0678
		(0.88)		(1.05)		(0.38)
Treatment(II)*Round		-0.0257		-0.0256		-0.0125
		(-0.20)		(-0.17)		(-0.07)
Treatment(III)*Round		0.126		0.0704		0.182
		(0.87)		(0.40)		(1.00)
Constant	0.719	-0.437	0.179	-1.878	2.322	2.728
	(0.65)	(-0.32)	(0.14)	(-1.34)	(1.12)	(0.94)
Demographic Controls	YES	YES	YES	YES	YES	YES
Observations	2680	2680	1672	1672	1008	1008
Individuals	335	335	209	209	126	126
Groups(Clusters)	87	87	83	83	79	79

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Notes: OLS random effects estimation. z-statistics in parentheses. Robust standard errors, clustered at group level. Estimates for the pooled sample. Included demographic controls: age, sex, risk aversion, correct answer to control question, reading-time, CRT-score, and working memory score.

Table 5 contains results from six different regressions. Specifications (1) and (2) use pooled data from the HC. We find no evidence that changes in behavior depend on treatment assignment. The significant negative coefficient of the *Round* variable captures a general decline in contributions. The effects for the two main variables of interest (*Above Group Average in t-1* and *Below Group Average in t-1*) point towards the presence of conditional cooperation. Subjects contributing more than the group average, decrease their contributions significantly in the subsequent round. Similarly, subjects who contribute less than the group average increase their contributions significantly in the following round. In line with Ashley et al. (2010), the coefficients of both variables differ in their strength. The fact that subjects react more strongly to lower contributions of their group members could be one important factor shaping the decline in average contributions across rounds. In specification (2) we test whether time pressure affects subjects in their ability to condition their behavior on the choices of the other group members. We capture these effects by interacting the treatment dummies with the main variables of interest. The interaction terms provide weak evidence contradicting our hypothesis that time pressure would decrease subjects' responsiveness to the choices of other group members. Instead, time pressure causes subjects to reduce their own contributions more strongly when observing lower average contributions by their group members. On the other hand, subjects under time pressure who contribute less than their group members do not increase their contributions in the following round as much as unconstrained subjects. This second interaction effect is, however, insignificant for the pooled sample. To confidently interpret changes in contribution behavior as conditional cooperation, subjects should display low levels of confusion regarding the underlying incentive structure.[20] Therefore, in specifications (3) - (6) we provide separate estimates based on subjects' confusion status by once more splitting the sample into two 'bins' according to behavior in the CC.[21] Specifications (3) and (5) show that both confused and unconfused subjects condition their own behavior on the past choices of their group members. A comparison of specifications (4) and (6) reveals that time pressure selectively affects the strategic behavior of unconfused subjects. As for the pooled sample, subjects under time pressure react more strongly to negative experiences with their group members. Time pressure, however, does not lead to an overall increase in conditional cooperation. Treated subjects are also more prone to exploit the higher cooperation levels of their group members by not increasing their own contributions. Especially this second observation contradicts an intuitive predisposition towards cooperative behavior.

Result 7: ***We find no evidence that subjects under time pressure are less able to condition their behavior on that of other subjects. They are more likely to reduce their contributions upon a negative experience with their group members, while they are less likely to increase their contributions after a positive experience. These interaction effects are only present among unconfused subjects.***

4 Discussion and Conclusion

In this paper we pursued two objectives. First, in light of some conflicting evidence (Fiedler et al., 2013a; Tinghög et al., 2013; Verkoeijen and Bouwmeester, 2014; Martinsson et al., 2014; Capraro and Cococcioni, 2016; Myrseth and Wollbrant, 2017; Bouwmeester et al., 2017; Kocher et al., 2017), we reinvestigate the initial claim that a more intuitive cognitive style causes subjects

[20] One plausible alternative explanation why confused subjects might condition their behavior on the choices of others could be that they see these choices as containing an informative signal about the game form (Burton-Chellew and West, 2013).

[21] To account for learning, we classify a subject as confused if his contributions across the nine rounds of the CC are above those of the average subject. Results are similar if we classify subjects according to their behavior in the final round.

in a linear PGG to behave more cooperatively (Rand et al., 2012, 2014). Our findings from the HC condition confirm that subjects' cognitive style indeed affects their choices. However, the direction of the effect goes in the opposite direction of previous findings (Rand et al., 2012, 2014; Rand, 2016). In our experiment, time pressure significantly increases the share of zero contributions and weakly decreases average contributions. This speaks against the conclusion that an intuitive cognitive style generally favors cooperative behavior. There are several subtle design differences between our experiment and the original studies by Rand et al. (2012, 2014). Our experiment draws on student subjects, while subjects in Rand et al. (2012) were recruited via Amazon Mechanical Turk and are thus more diverse in background. We use higher stakes and the time limit in our experiment is slightly stronger. None of these differences, however, should affect the direction of the treatment effect, if intuition was linked to cooperation in a general way as suggested by the social heuristics hypothesis (SHH) (Rand et al., 2014). This conclusion survives several additional robustness checks. Variations of the time pressure protocol and an extension to repeat interaction both show that time pressure either significantly reduces contributions or does not affect contributions, but never significantly increases contributions in these test, as the SHH would predict. Evidence from repeat interactions also indicates that time pressure has a limited effect on subjects' ability to condition their behaviour on previous choices of other group members.

The observations in this paper speak not only to the literature on cognitive style and cooperation. It also speaks to a literature that is interested in the link between self-control and cooperation.[22] This literature strongly suggests that self-control could be essential for cooperation (Martinsson et al., 2014; Kocher et al., 2017): There is a positive association between higher patience measured in time discounting tasks - a sign of higher self-control - and cooperative behaviour in lab and field settings (Curry et al., 2008; Burks et al., 2009; Fehr and Leibbrandt, 2011). There is also a direct correlation between self-control as a measurable trait and cooperative behaviour, in particular when a self-control conflict is made salient (Martinsson et al., 2014; Myrseth et al., 2015; Kocher et al., 2017). These findings are closely aligned with results 1 and 2 of our paper and suggest that heightened self-control and more deliberative decision-making could have similar effects on cooperative behavior. The findings from this literature might also be providing a novel and additional explanation for the decreasing contribution trend in repeated PGG games. If it becomes harder to exert self-control over time - as suggested by parts of the literature (Achtziger et al., 2015) - and self-control is required for cooperation, then a declining trend might be driven by a depletion of self-control resources.

The second objective of this study was to investigate whether cognitive failure (confusion) is the source of an important confound, when time pressure is used to investigate the link between a more intuitive cognitive style and cooperation in PGG. Regarding the role of cognitive failure our results are twofold. Contrary to previous concerns, voiced in the context of response time studies (Recalde et al., 2014), we find no evidence that forcing subjects to decide quickly increases confusion. More precisely, the results in Recalde et al. (2014) show that faster subjects are more prone to making an error in a non-linear public goods game (presumably because it takes some time to calculate the optimal interior contribution level), while our results show that exogenously

[22]We are grateful to an anonymous referee for pointing out this connection.

speeding up decision makers does not induce more errors in a standard public good game. This is good news for studies that use time pressure to induce more intuitive decision making in the PGG, as this means that results such as those of Rand et al. (2012, 2014) are not merely an artifact of inducing more confusion. Of course, our finding does not rule out that time pressure might increase confusion or reduce decision quality in other, more complex decision tasks or when the trade-off between decision speed and decision quality is made more salient as in the PEI or PRC protocols.

Even within the comparably simple setting of a linear PGG, behavior in the CC closely replicates earlier studies (Houser and Kurzban, 2002) in pointing towards a substantial presence of confusion; approximately 50 percent of subjects in our experiment contribute in a task where contributing decreases their own earnings without providing an efficiency gain for other participants. Based on this behavioral measure of cognitive failure we show that confusion status is an important driver of individual heterogeneity in contribution behavior. It affects the level of contributions and the distribution of contributions. Most importantly for the objective of this study, confusion status is the source of a heterogeneous treatment effect: time pressure selectively affects unconfused subjects by reducing their contributions. This is partly in line with findings in Strømland et al. (2016), who also identify confusion as a key moderator. This heterogeneous treatment effect could be one key explanation for why our reinvestigation of the TP treatment effect, along with those of others, fails to replicate and possibly contradicts the earlier findings of the role of cognitive style on cooperation. The presence of this moderating factor complicates the comparison of time pressure effects across different experiments, especially if the extent of cognitive failure varies between different experimental populations. This could be the case when drawing on non-student samples (Belot et al., 2015) or using different sets of instructions (Ferraro and Vossler, 2010). More generally, our results highlight the importance of cognitive failure as an understudied source of contribution heterogeneity in the PGG and as a potential moderator that can affect the internal validity of experimental results. The methods we have applied to identify confusion are easily replicable and adaptable to other designs and might provide for a more comprehensive robustness check than non-behavioral measures of confusion.

References

A. Achtziger, C. Als-Ferrer, and A. K. Wagner. Money, depletion, and prosociality in the dictator game. *Journal of Neuroscience, Psychology, and Economics*, 8(1):1, 2015.

J. Andreoni. Why free ride? Strategies and learning in public goods experiments. *Journal of Public Economics*, 37(3):291–304, 1988.

J. Andreoni. Cooperation in public-goods experiments: Kindness or confusion? *The American Economic Review*, 85(4):891–904, 1995.

J. D. Angrist, G. W. Imbens, and D. B. Rubin. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455, 1996.

R. Ashley, S. Ball, and C. Eckel. Motives for Giving: A Reanalysis of Two Classic Public Goods Experiments. *Southern Economic Journal*, 77(1):15–26, 2010.

R.-C. Bayer, E. Renner, and R. Sausgruber. Confusion and learning in the voluntary contributions game. *Experimental Economics*, 16(4):1–19, 2013.

M. Belot, R. Duch, and L. Miller. A comprehensive comparison of students and non-students in classic experimental games. *Journal of Economic Behavior & Organization*, 113:26–33, 2015.

H. S. Bloom. Accounting for no-shows in experimental evaluation designs. *Evaluation Review*, 8(2):225–246, 1984.

S. Bouwmeester, P. P. Verkoeijen, B. Aczel, F. Barbosa, L. Bègue, P. Brañas-Garza, T. G. Chmura, G. Cornelissen, F. S. Døssing, A. M. Espín, et al. Registered replication report: Rand, greene, and nowak (2012). *Perspectives on Psychological Science*, 12(3):527–542, 2017.

S. V. Burks, J. P. Carpenter, L. Goette, and A. Rustichini. Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences*, 106(19):7745–7750, 2009.

M. N. Burton-Chellew and S. A. West. Prosocial preferences do not explain human cooperation in public-goods games. *Proceedings of the National Academy of Sciences*, 110(1):216–221, 2013.

M. N. Burton-Chellew, C. El Mouden, and S. A. West. Conditional cooperation and confusion in public-goods experiments. *Proceedings of the National Academy of Sciences*, 113(5):1291–1296, 2016.

A. W. Cappelen, U. H. Nielsen, B. Tungodden, J.-R. Tyran, and E. Wengström. Fairness is intuitive. *Experimental Economics*, 19(4):727–740, 2016.

D. Cappelletti, W. Güth, and M. Ploner. Being of two minds: Ultimatum offers under cognitive constraints. *Journal of Economic Psychology*, 32(6):940–950, 2011.

V. Capraro and G. Cococcioni. Rethinking spontaneous giving: Extreme time pressure and ego-depletion favor self-regarding reactions. *Nature Scientific reports*, 6:27219, 2016.

V. Capraro, J. J. Jordan, and D. G. Rand. Heuristics guide the implementation of social preferences in one-shot prisoner's dilemma experiments. *Nature Scientific Reports*, 4, 2014.

J. Carpenter, M. Graham, and J. Wolf. Cognitive ability and strategic sophistication. *Games and Economic Behavior*, 80:115–130, 2013.

A. Chaudhuri. Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics*, 14(1):47–83, 2011.

E. Chou, M. McConnell, R. Nagel, and C. R. Plott. 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, 2009.

J. Cone and D. G. Rand. Time pressure increases cooperation in competitively framed social dilemmas: a successful replication. *PloS one*, 9(12), 2014.

B. Corgnet, A. M. Espín, and R. Hernán-González. The cognitive basis of social behavior: Cognitive reflection overrides antisocial but not always prosocial motives. Working Paper, 2015.

R. Croson, E. Fatas, and T. Neugebauer. Reciprocity, matching and conditional cooperation in two public goods games. *Economics Letters*, 87(1):95–101, 2005.

R. T. Croson. Theories of commitment, altruism and reciprocity: Evidence from linear public goods games. *Economic Inquiry*, 45(2):199–216, 2007.

O. S. Curry, M. E. Price, and J. G. Price. Patience is a virtue: Cooperative people have lower discount rates. *Personality and individual differences*, 44(3):780–785, 2008.

A. Dreber, D. Fudenberg, D. K. Levine, and D. G. Rand. Altruism and Self-Control. *Discussion Paper(Available at SSRN: 2477454)*, 2014.

S. Duffy and J. Smith. Cognitive load in the multi-player prisoner's dilemma game: Are there brains in games? *Journal of Behavioral and Experimental Economics*, 51:47–56, 2014.

J. S. B. T. Evans. Dual-processing accounts of reasoning, judgment, and social cognition. *Annu. Rev. Psychol.*, 59:255–278, 2008.

E. Fehr and A. Leibbrandt. A field study on cooperativeness and impatience in the tragedy of the commons. *Journal of Public Economics*, 95(9-10):1144–1155, 2011.

P. J. Ferraro and C. A. Vossler. The source and significance of confusion in public goods experiments. *The BE Journal of Economic Analysis & Policy*, 10(1):1935–1682, 2010.

S. Fiedler, A. Glöckner, A. Nicklisch, and S. Dickert. Social value orientation and information search in social dilemmas: An eye-tracking analysis. *Organizational behavior and human decision processes*, 120(2):272–284, 2013a.

S. Fiedler, A. Glöckner, A. Nicklisch, and S. Dickert. Social value orientation and information search in social dilemmas: An eye-tracking analysis. *Organizational Behavior and Human Decision Processes*, 120(2):272–284, 2013b.

U. Fischbacher and S. Gächter. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review*, 100(1):541–556, 2010.

U. Fischbacher, S. Gächter, and E. Fehr. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, 71(3):397–404, 2001.

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

G. R. Fréchette. Session-effects in the laboratory. *Experimental Economics*, 15(3):485–498, 2012.

S. Frederick. Cognitive reflection and decision making. *The Journal of Economic Perspectives*, 19(4):25–42, 2005.

S. Gächter. Human behaviour: A cooperative instinct. *Nature*, 489(7416):374–375, 2012.

M. Gärtner and A. Sandberg. Is there an omission effect in prosocial behavior? a laboratory experiment on passive vs. active generosity. *PloS one*, 12(3):e0172496, 2017.

B. Greiner. An online recruitment system for economic experiments. *MPRA Working Paper Series in Economics 13513*, 2004.

Z. Grossman, J. J. Van der Weele, and A. Andrijevik. A test of dual-process reasoning in charitable giving. *Working Paper (Available at SSRN: 2520585)*, 2014.

K. E. Hauge, K. A. Brekke, L.-O. Johansson, O. Johansson-Stenman, and H. Svedsäter. Keeping others in our mind or in our heart? distribution games under cognitive load. *Experimental Economics*, 19(3):562–576, 2016.

D. Houser and R. Kurzban. Revisiting kindness and confusion in public goods experiments. *The American Economic Review*, 92(4):1062–1069, 2002.

G. W. Imbens and J. D. Angrist. Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475, 1994.

D. Kahneman. *Thinking, fast and slow.* Farrar, Straus and Giroux, 2011.

M. G. Kocher and M. Sutter. Time is money – time pressure, incentives, and the quality of decision-making. *Journal of Economic Behavior & Organization*, 61(3):375–392, 2006.

M. G. Kocher, P. Martinsson, K. O. R. Myrseth, and C. E. Wollbrant. Strong, bold, and kind: Self-control and cooperation in social dilemmas. *Experimental Economics*, 20:44–69, 2017.

I. Krajbich, B. Bartling, T. Hare, and E. Fehr. Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nature Communications*, 6, 2015.

J. O. Ledyard. Public goods: A survey of experimental research. In J. Kagel and A. Roth, editors, *Handbook of experimental economics*. Princeton University Press, Princeton, 1995.

G. Loewenstein and T. O' Donoghue. Animal spirits: Affective and deliberative processes in economic behavior. *Working Paper (Available at SSRN 539843)*, 2007.

J. Lohse. Smart or selfish–when smart guys finish nice. *Journal of Behavioral and Experimental Economics*, 64:28–40, 2016.

J. Lohse, T. Goeschl, and J. H. Diederich. Giving is a question of time: response times and contributions to an environmental public good. *Environmental and Resource Economics*, 67 (3):455–477, 2017.

P. Martinsson, K. O. R. Myrseth, and C. Wollbrant. Social dilemmas: When self-control benefits cooperation. *Journal of Economic Psychology*, 45:213–236, 2014.

A. L. Merkel and J. Lohse. Is fairness intuitive? an experiment accounting for subjective utility differences under time pressure. *Experimental Economics*, pages 1–27, 2018.

M. Milinski and C. Wedekind. Working memory constrains human cooperation in the prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 95(23):13755–13758, 1998.

K. O. R. Myrseth and C. E. Wollbrant. Cognitive foundations of cooperation revisited: Commentary on Rand et al.(2012, 2014). *Journal of Behavioral and Experimental Economics*, 69: 133–138, 2017.

K. O. R. Myrseth, G. Riener, and C. E. Wollbrant. Tangible temptation in the social dilemma: Cash, cooperation, and self-control. *Journal of Neuroscience, Psychology, and Economics*, 8 (2):61, 2015.

U. H. Nielsen, J.-R. Tyran, and E. Wengström. Second thoughts on free riding. *Economics Letters*, 122(2):136–139, 2014.

M. Piovesan and E. Wengström. Fast or fair? A study of response times. *Economics Letters*, 105(2):193–196, 2009.

D. G. Rand, J. D. Greene, and M. A. Nowak. Spontaneous giving and calculated greed. *Nature*, 489(7416):427–430, 2012.

D. G. Rand. Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychological Science, Forthcoming*, 2016.

D. G. Rand, A. Peysakhovich, G. T. Kraft-Todd, G. E. Newman, O. Wurzbacher, M. A. Nowak, and J. D. Greene. Social heuristics shape intuitive cooperation. *Nature Communications*, 5, 2014.

D. G. Rand, G. E. Newman, and O. M. Wurzbacher. Social context and the dynamics of cooperative choice. *Journal of behavioral decision making*, 28(2):159–166, 2015.

M. P. Recalde, A. Riedl, and L. Vesterlund. Error prone inference from response time: The case of intuitive generosity. *CESifo Working Paper Series*, 4987, 2014.

S. G. Roch, J. A. Lane, C. D. Samuelson, S. T. Allison, and J. L. Dent. Cognitive load and the equality heuristic: A two-stage model of resource overconsumption in small groups. *Organizational Behavior and Human Decision Processes*, 83(2):185–212, 2000.

A. Rubinstein. Instinctive and cognitive reasoning: A study of response times. *The Economic Journal*, 117(523):1243–1259, 2007.

A. Rubinstein. Response time and decision making: An experimental study. *Judgment and Decision Making*, 8(5):540–551, 2013.

A. Schotter. What's so informative about choice? In A. Schotter and A. Caplin, editors, *The foundations of positive and normative economics: a handbook*, pages 70–94. Oxford University Press, 2008.

A. Schotter and I. Trevino. Is response time predictive of choice? An experimental study of threshold strategies. *WZB Discussion Paper No. SP II 2014-305*, 2014.

L. Spiliopoulos and A. Ortmann. The bcd of response time analysis in experimental economics. *Experimental Economics*, pages 1–51, 2016.

E. Strømland, S. Tjøtta, and G. Torsvik. Cooperating, fast and slow: Testing the social heuristics hypothesis. *CESifo Working Paper Series*, 2016.

G. Tinghög, D. Andersson, C. Bonn, H. Böttiger, C. Josephson, G. Lundgren, D. Västfjäll, M. Kirchler, and M. Johannesson. Intuition and cooperation reconsidered. *Nature*, 498(7452): E1–E2, 2013.

S. T. Trautmann. Risk taking under stress: the role (s) of self-selection. a comment on buckert et al.(2014). *Frontiers in neuroscience*, 8:197, 2014.

P. Ubeda. The consistency of fairness rules: An experimental study. *Journal of Economic Psychology*, 41(0):88–100, 2014.

P. P. Verkoeijen and S. Bouwmeester. Does intuition cause cooperation? *PloS one*, 9(5), 2014.

D. Wechsler. Wechsler Adult Intelligence Scale. *Psycholgical Corporation*, 1955.

Appendix

Response Time Distribution

Table 6 summarizes the distribution of response times in the one-shot public good game across the four HC.

Table 6: Response times: human condition across the different treatment and order conditions

	Normal Order		**Reverse Order**	
Percentile	(HC-BL)	(HC-TP)	(HC-BL)	(HC-TP)
1%	6	4	4	2
5%	7	4	4	3
10%	8	5	5	3
25%	11.5	6	6.5	3
50%	16	7	10	4
75%	23	8	13	6
90%	33	10	18	7
95%	56	17	20	8
99%	109	25	40	26

Notes: Response time percentiles for the one-shot public good game across the different order conditions.

Further Robustness Checks For Online Appendix

In tobit regression models (1) and (2) of Table 7 we analyze the effect of treatment assignment on contributions. Relative to observations from the normal order condition without time pressure, subjects under time pressure contribute less under normal task order (*Treatment(I): Normal Order + Time Pressure*). This effect is weakly significant at the 10 percent level. Being assigned to the reverse order condition (*Treatment(II): Reverse Order*) reduces contributions, but not significantly. Applying time pressure in the reverse order condition (*Treatment(III): Reverse Order + Time Pressure*) significantly reduces contributions relative to subjects in the normal order condition without time pressure, but not relative to subjects in the reverse order condition (Wald Test: p=0.5545). These results continue to hold when controlling for the same demographic variables as in Rand et al. (2012) (age, sex, ability to answer comprehension question correctly), a survey measure of risk aversion, and several psychometric variables (time spent on the instruction screen, CRT-score, and working memory test score). Probit regression models (3) and (4) estimate the effect of treatment assignment on the propensity to contribute zero tokens. Again time pressure significantly increases free-riding without further control variables and when using the same covariates as for contribution behavior. Contributions in the additional time pressure protocols (*Treatment(IV): Normal Order + PEI* and *Treatment(VI): Normal Order + PRC*) are neither significantly larger nor significantly smaller compared to a condition without time pressure. The effects on free-riding behaviour are also insignificant.

Table 7: Effects of treatment assignment on contributions in HC

	(1) Contributions	(2) Contributions	(3) Free-Riding	(4) Free-Riding
Treatment(I): Normal Order + Time Pressure (1=Yes)	-3.230*	-3.340*	0.412**	0.430**
	(-1.80)	(-1.88)	(2.01)	(2.02)
Treatment(II): Reverse Order (1=Yes)	-3.361	-3.098	0.343	0.408*
	(-1.62)	(-1.54)	(1.45)	(1.65)
Treatment(III): Reverse Order + Time Pressure (1=Yes)	-4.779**	-5.474***	0.594***	0.658***
	(-2.22)	(-2.63)	(2.60)	(2.73)
Treatment(IV): Normal Order + PEI (1=Yes)	-0.897	-1.821	0.182	0.290
	(-0.42)	(-0.77)	(0.71)	(0.99)
Treatment(VI): Normal Order + PRC (1=Yes)	0.580	-0.737	0.137	0.316
	(0.28)	(-0.33)	(0.54)	(1.12)
Age (Years)		0.0512		-0.0294
		(0.24)		(-1.12)
Sex (1=Male)		3.309**		-0.0501
		(2.57)		(-0.34)
Cognitive Reflection Score (0-3)		1.732***		0.0584
		(3.26)		(0.90)
Risk Aversion (1-11)		0.0737		-0.00101
		(0.30)		(-0.03)
Unconfused (1=Yes)		-3.682***		0.377**
		(-2.91)		(2.57)
Readingtime Instructions (Log10 Sec.)		3.106**		-0.234
		(2.05)		(-1.31)
Working Memory Score (0-1)		3.831		-0.527
		(0.96)		(-1.21)
Constant	12.79****	-3.266	-1.173****	0.159
	(10.12)	(-0.43)	(-7.50)	(0.17)
Observations	464	451	464	451

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Notes: Specifications (1)-(2): Tobit estimation to account for censoring from below (0) and above (20). Specifications (3)-(4): Probit estimation. t-statistics in parentheses. Robust standard errors. Estimates for the full sample.

IV-Regression results: Robustness checks contributions

Table 8 demonstrates that the results for contributions reported in table 2 are robust to the following alternative specifications: (1) using only obsercations from the normal order condition, (2) OLS instead of Tobit, (3, 4) using a dummy variable for fast decisions (response times either <= 5 or <= 7 seconds) instead of a continuous response time variable. Specifications (5)-(8) arrive at the same conclusion when controlling for a set of additional demographic and psychometric attributes.

Table 8: Effects of treatment assignment on contributions - Robustness Checks

	(1) Contrib	(2) Contrib	(3) Contrib	(4) Contrib	(5) Contrib	(6) Contrib	(7) Contrib	(8) Contrib
Response Time (Log)	3.789*	1.908**			3.957*	2.184***		
	(1.79)	(2.27)			(1.75)	(2.60)		
Below 7 (1=Yes)			-5.298**				-6.453***	
			(-2.20)				(-2.60)	
Below 5 (1=Yes)				-6.049**				-7.123**
				(-1.96)				(-2.31)
Age (Years)					-0.323	-0.155	-0.173	-0.164
					(-0.86)	(-0.99)	(-0.62)	(-0.59)
Sex (1=Male)					3.523**	1.964**	3.593**	3.585**
					(1.99)	(2.46)	(2.50)	(2.49)
Risk Aversion (1-11)					0.156	0.0858	0.232	0.189
					(0.42)	(0.51)	(0.77)	(0.63)
Unconfused (1=Yes)					-1.568	-1.292	-2.111	-1.901
					(-0.88)	(-1.57)	(-1.43)	(-1.27)
Cognitive Reflection Score (0-3)					1.606*	1.236***	1.831***	1.798***
					(1.82)	(3.30)	(2.74)	(2.69)
Readingtime Instructions (Log10 Sec.)					2.296	1.616*	3.676**	3.722**
					(1.13)	(1.76)	(2.20)	(2.25)
Working Memory Score					1.632	1.208	4.519	2.832
					(0.26)	(0.48)	(0.94)	(0.61)
Constant	2.075	5.582***	12.55****	11.69****	-3.713	-0.303	-2.164	-2.745
	(0.40)	(2.95)	(9.81)	(11.47)	(-0.32)	(-0.06)	(-0.24)	(-0.31)
Observations	220	348	348	348	215	335	335	335

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

IV-Regression results: Robustness checks free-riding

Table 9 demonstrates that the results for freeriding reported in table 2 are robust to the following alternative specifications: (1) using only observations from the normal order condition, (2) OLS instead of Probit, (3, 4) using a dummy variable for fast decisions (response times either <= 5 or <= 7 seconds) instead of a continuous response time variable. Specifications (5)-(8) arrive at the same conclusion when controlling for a set of additional demographic and psychometric attributes.

Table 9: Effects of treatment assignment on free-riding - Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Free	Free	Free	Free	Free	Free	Free	Free
logtime1	-0.467**	-0.121***			-0.540**	-0.130***		
	(-2.04)	(-2.74)			(-2.05)	(-2.79)		
below7			0.664***				0.760***	
			(2.60)				(2.69)	
below5				0.693**				0.762**
				(2.35)				(2.48)
Age (Years)					-0.0314	-0.00266	-0.0163	-0.0155
					(-0.70)	(-0.30)	(-0.50)	(-0.49)
Sex (1=Male)					-0.329	-0.0353	-0.163	-0.160
					(-1.49)	(-0.80)	(-0.97)	(-0.96)
Risk Aversion (1-11)					-0.00456	-0.000708	-0.00673	-0.00140
					(-0.10)	(-0.08)	(-0.19)	(-0.04)
Unconfused (1=Yes)					0.427*	0.0620	0.234	0.208
					(1.92)	(1.41)	(1.38)	(1.24)
Cognitive Reflection Score (0-3)					0.106	0.0105	0.0627	0.0663
					(0.99)	(0.52)	(0.79)	(0.84)
Readingtime Instructions (Log10 Sec.)					0.269	-0.0167	-0.117	-0.116
					(0.94)	(-0.30)	(-0.56)	(-0.57)
Working Memory Score (0-1)					-0.390	-0.137	-0.748	-0.514
					(-0.63)	(-1.02)	(-1.49)	(-1.09)
Constant	0.165	0.464****	-1.191****	-1.039****	0.00339	0.624*	-0.312	-0.267
	(0.29)	(4.50)	(-8.02)	(-9.63)	(0.00)	(1.96)	(-0.29)	(-0.25)
Observations	220	348	348	348	215	335	335	335

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

IV-Regression results: Robustness Including PRC and PEI

Table 10 demonstrates that the results for contributions reported in table 2 are robust to including observations from the PEI and PRC treatments.

Table 10: Effects of treatment assignment

	(1)	(2)	(3)	(4)
	Contributions	Contributions	Free-Riding	Free-Riding
Panel A: Second Stage (DV: Contributions)	**HC**	**HC**	**HC**	**HC**
Response Time (Log Seconds)	3.105**	4.043***	-0.408**	-0.487***
	(2.05)	(2.66)	(-2.49)	(-2.80)
Age (Years)		-0.00267		-0.0261
		(-0.01)		(-0.98)
Sex (1=Male)		3.351***		-0.0693
		(2.64)		(-0.47)
Risk Aversion (1-11)		0.0552		0.00328
		(0.22)		(0.11)
Unconfused (1=Yes)		-4.006***		0.397***
		(-3.21)		(2.69)
Cognitive Reflection Score (0-3)		1.765****		0.0522
		(3.34)		(0.77)
Readingtime Instructions (Log10 Sec.)		2.441*		-0.129
		(1.65)		(-0.70)
Working Memory Score (0-1)		7.284**		-0.806**
		(2.05)		(-2.07)
Constant	4.165	-12.27	-0.0284	1.221
	(1.25)	(-1.48)	(-0.08)	(1.26)
Panel B: First Stage (DV: Response Time)				
Treatment(I): Normal Order + Time Pressure (1=Yes)	-0.856****	-0.817****	-0.855****	-0.817****
	(-12.53)	(-11.89)	(-13.55)	(-13.02)
Treatment(II): Reverse Order (1=Yes)	-0.582****	-0.611****	-0.579****	-0.611****
	(-6.76)	(-7.30)	(-7.84)	(-8.34)
Treatment(III): Reverse Order + Time Pressure (1=Yes)	-1.326****	-1.328****	-1.325****	-1.328****
	(-17.02)	(-15.98)	(-17.94)	(-17.90)
Treatment(IV): Normal Order + PEI (1=Yes)	-0.919****	-0.837****	-0.923****	-0.840****
	(-11.82)	(-9.19)	(-11.94)	(-10.00)
Treatment(VI): Normal Order + PRC (1=Yes)	-0.700****	-0.672****	-0.706****	-0.678****
	(-8.75)	(-7.39)	(-9.36)	(-8.31)
Age (Years)		0.00763		0.00756
		(0.76)		(0.99)
Sex (1=Male)		-0.0294		-0.0297
		(-0.68)		(-0.66)
Risk Aversion (1-11)		0.00599		0.00602
		(0.62)		(0.66)
Unconfused (1=Yes)		0.0436		0.0432
		(0.97)		(0.98)
Cognitive Reflection Score (0-3)		-0.0319		-0.0322
		(-1.55)		(-1.58)
Readingtime Instructions (Log10 Sec.)		0.224****		0.225****
		(3.73)		(4.05)
Working Memory Score (0-1)		-0.308**		-0.302**
		(-2.30)		(-2.29)
Constant	2.828****	2.039****	2.829****	2.037****
	(48.61)	(7.14)	(62.79)	(7.35)
Observations	464	451	464	451

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Notes: Specifications (1) and (2): Tobit-IV maximum likelihood estimation to account for censoring from below (0) and above (20). Specifications (3)-(4): Probit-IV maximum likelihood estimation. t-statistics in parentheses. Robust standard errors. Estimates for the pooled sample including the PEI and PRC treatments.

IV-Regression results: Robustness heterogeneous treatment effects

Table 11 demonstrates that the results for contributions reported in table 3 are robust to including additional observations from the PEI and PRC treatments.

Table 11: Instrumental variable estimates of the effects of fast decision making separated by confusion status containing additional time pressure treatments

	Unconfused		Confused	
	(1)	(2)	(3)	(4)
	Contributions	Contributions	Contributions	Contributions
Response Time (Log10 Seconds)	7.957**		2.483	
	(2.33)		(1.62)	
Response Time ≤ 7 Sec. (1=Yes)		-12.24**		-4.153
		(-2.48)		(-1.60)
Age (Years)	-0.347	-0.257	0.114	0.117
	(-0.66)	(-0.49)	(0.63)	(0.63)
Sex (1=Male)	3.587	3.907	4.108****	4.287****
	(1.29)	(1.40)	(3.31)	(3.39)
Risk Aversion (1-11)	-0.311	-0.439	0.345	0.373
	(-0.54)	(-0.77)	(1.50)	(1.56)
Cognitive Reflection Score (0-3)	2.450*	2.207*	1.986****	1.936****
	(1.89)	(1.72)	(3.99)	(3.82)
Readingtime Instructions (Log10 Sec.)	6.102*	6.635**	-0.110	-0.0791
	(1.76)	(1.96)	(-0.08)	(-0.06)
Working Memory Score (0-1)	11.54	13.67	4.678	5.228*
	(1.28)	(1.52)	(1.56)	(1.65)
Constant	-32.24	-13.18	-3.950	2.735
	(-1.62)	(-0.72)	(-0.53)	(0.41)
Observations	213	213	238	238

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Notes: Tobit-IV maximum likelihood estimation to account for censoring from below (0) and above (20). Specifications (1)-(2): subjects classified as unconfused. Specifications (3)-(4): subjects classified as confused. t-statistics in parentheses. Robust standard errors. Estimates for the pooled sample.

Results for pooled sample of repeated PGG

Regressions (1) and (2) display results from the HC, using individual contributions in round t as the dependent variable. Regressions (3) and (4) similarly model contribution behavior in the CC. In each specification subjects from the normal order condition serve as the left-out baseline category, against which we compare behavior in the other three randomly assigned treatment conditions.

Table 12: Repeated decisions: Decay of contributions HC and CC

	Human Condition (HC)		Computer Condition (CC)	
	(1)	(2)	(3)	(4)
	Contributions	Contributions	Contributions	Contributions
Treatment(I): Normal Order + Time Pressure (1=Yes)	-2.071**	-1.969*	0.035	-0.806
	(-2.54)	(-1.84)	(0.06)	(-0.85)
Treatment(II): Reverse Order (1=Yes)	-3.003**	-2.840*	0.494	0.186
	(-2.52)	(-1.94)	(0.74)	(0.16)
Treatment(III): Reverse Order + Time Pressure (1=Yes)	-2.300*	-3.017**	1.195*	0.958
	(-1.85)	(-2.20)	(1.77)	(0.86)
Round (1-9)	-0.658****	-0.664****	-0.402****	-0.484****
	(-9.68)	(-5.22)	(-8.26)	(-5.81)
Last Round (1=Yes)	-0.374	-0.674	0.294	0.643*
	(-1.00)	(-0.90)	(1.34)	(1.69)
Treatment(I)*Round		-0.043		0.191
		(-0.25)		(1.60)
Treatment(II)*Round		-0.024		0.062
		(-0.12)		(0.44)
Treatment(III)*Round		0.140		0.048
		(0.70)		(0.33)
Treatment(I)*Last Round		1.023		-1.036**
		(1.09)		(-1.99)
Treatment(II)*Last Round		-0.369		-0.031
		(-0.35)		(-0.05)
Treatment(III)*Last Round		0.176		-0.016
		(0.15)		(-0.02)
Constant	7.900**	7.963**	3.137	3.511
	(2.05)	(2.05)	(1.04)	(1.16)
Demographic Controls	YES	YES	YES	YES
Observations	3015	3015	3015	3015
Individuals	335	335	335	335
Groups(Clusters)	87	335	87	335
R^2	0.15	0.15	0.07	0.07
Prob > Chi^2	<0.001	<0.001	<0.001	<0.001

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Notes: OLS random effects estimation. z-statistics in parentheses. Robust standard errors, clustered at group level (HC) or individual level (CC). Estimates for the pooled sample. Included demographic controls: age, sex, risk aversion, correct answer to control question, reading-time, CRT-score, and working memory score.

Regressions (1) and (2) show that applying time pressure significantly reduces contributions in the normal order condition of the HC. Furthermore, exogenously increasing subjects experience by assigning them to the reverse order condition reduces contributions: the coefficients of both the *Treatment(II)* and *Treatment(III)* dummies are negative and significant. However, applying time pressure under reverse order does not further reduce contributions.[23] In the CC there is

[23]This can be verified by comparing the size of the coefficients of *Treatment(II)* and *Treatment(III)*. A Wald

little evidence that the different treatment conditions affect contribution behavior. Only inexperienced subjects deciding under time pressure (*Treatment(III): Reverse Order + Time Pressure*) display marginally increased contribution levels. The decline in contributions is captured by the *Round* variable, which is negative and significant across all specifications - as typical for public good games. The decline is steeper in the HC than in the CC. To test for potential end-game effects we include an additional dummy variable indicating the last round, which is insignificant, both in the HC and CC. Finally, by including interaction terms in regressions (2) and (4) we analyze, if deciding under a time limit or deciding in the reverse order condition affects the decay of contributions. One plausible hypothesis would be that constraining deliberation via time pressure negatively affects the rate of learning, because subjects can invest lower cognitive efforts to understand the game form or the behavior of their group members. We find no support for this hypothesis in the HC. Despite constraining deliberation via time pressure or giving subjects additional experience in the reverse order conditions, contributions decline at comparable speeds. In the CC there is weak evidence that subjects under time pressure converge slower in the early rounds, but display faster convergence in the last round.

test fails to reject the hypothesis that they are the same (Chi^2: χ^2=0.22 p=0.6383). The same conclusion can be drawn from an alternative specification including interaction terms between a time pressure and an order dummy or by estimating separate regressions for observations from the normal and reverse order condition.

Robustness repeated game evidence HC

Table 13 contains several robustness checks to the results discussed in table 12 regarding repeated game cooperation in the HC. Specifications (1) and (2) include observations from the PEI and PRC treatments. While the effects of the main time pressure treatment remains unchanged these additional treatments have no significant effect on contributions reducing them only by a smaller amount. Specifications (3) and (4) demonstrate that previous results are robust to estimating the model as a Tobit Model and specifications (5) and (6) cluster all standard errors at the individual level instead of at the group level.

Table 13: Repeated decisions: Decay of contributions HC and CC

	(1) Contrib HC	(2) Contrib HC	(3) Contrib HC	(4) Contrib HC	(5) Contrib HC	(6) Contrib HC
Treatment(I): Normal Order + Time Pressure (1=Yes)	-2.274***	-2.172**	-3.419**	-3.025*	-2.230***	-2.128**
	(-2.65)	(-1.99)	(-2.20)	(-1.73)	(-3.03)	(-2.08)
Treatment(II): Reverse Order (1=Yes)	-3.183***	-3.021**	-6.421****	-4.598**	-3.158****	-2.996**
	(-2.61)	(-2.03)	(-3.52)	(-2.23)	(-3.50)	(-2.38)
Treatment(III): Reverse Order + Time Pressure (1=Yes)	-2.514**	-3.231**	-3.862**	-4.472**	-2.514***	-3.231***
	(-1.97)	(-2.34)	(-2.11)	(-2.16)	(-2.70)	(-2.78)
Treatment(IV): PEI (1=Yes)	-0.309	-1.605				
	(-0.23)	(-1.23)				
Treatment(IV): PRC (1=Yes)	0.366	-0.461				
	(0.34)	(-0.37)				
Round (1-9)	-0.612****	-0.664****	-1.185****	-1.110****	-0.658****	-0.664****
	(-11.04)	(-5.23)	(-16.27)	(-8.76)	(-11.24)	(-6.31)
Last Round (1=Yes)	-0.147	-0.674	-0.794	-1.135	-0.374	-0.674
	(-0.47)	(-0.90)	(-1.31)	(-1.08)	(-1.14)	(-1.12)
Treatment(I)*Round		-0.0431		-0.116		-0.0431
		(-0.25)		(-0.65)		(-0.28)
Treatment(II)*Round		-0.0242		-0.358		-0.0242
		(-0.12)		(-1.63)		(-0.15)
Treatment(III)*Round		0.140		0.132		0.140
		(0.70)		(0.62)		(0.85)
Treatment(IV)*Round		0.240				
		(1.26)				
Treatment(V)*Round		0.133				
		(0.84)				
Treatment(I)*Last Round		1.023		1.719		1.023
		(1.09)		(1.17)		(1.20)
Treatment(II)*Last Round		-0.369		-1.407		-0.369
		(-0.35)		(-0.75)		(-0.40)
Treatment(III)*Last Round		0.176		-0.252		0.176
		(0.15)		(-0.14)		(0.19)
Treatment(IV)*Last Round		0.883				
		(0.85)				
Treatment(V)*Last Round		1.466				
		(1.33)				
Constant	7.632**	7.950**	7.355	7.000	7.247**	7.310**
	(2.20)	(2.26)	(0.91)	(0.86)	(2.03)	(2.04)
Demographic Controls	YES	YES	YES	YES	YES	YES
Observations	4059	4059	3015	3015	3015	3015
Individuals	451	451	335	335	335	335
Clusters(Groups)	116	116	-	-	335	335

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Notes: Included demographic controls: age, sex, risk aversion, correct answer to control question, reading-time, CRT-score, and working memory score.

Robustness of strategic cooperation results

Table 14 demonstrates that the results in table 5 are robust to categorizing the below and above variables based on tertiles.

Table 14: Repeated decisions: conditional cooperation

	HC: Full Sample		HC: Unconfused		HC: Confused	
	(1)	(2)	(3)	(4)	(5)	(6)
	Δ Contributions	Δ Contributions	Δ Contributions	Δ Contributions	Δ Contributions	Δ Contributions
Treatment(I): Normal Order + Time Pressure (1=Yes)	-0.0726	0.534	-0.0650	1.071	0.0419	-0.653
	(-0.36)	(0.82)	(-0.27)	(1.23)	(0.09)	(-0.55)
Treatment(II): Reverse Order (1=Yes)	-0.115	1.316*	0.166	2.117**	-0.814*	-0.161
	(-0.50)	(1.91)	(0.65)	(2.31)	(-1.67)	(-0.14)
Treatment(III): Reverse Order + Time Pressure (1=Yes)	-0.0149	-0.117	0.0287	0.145	-0.263	-0.903
	(-0.06)	(-0.18)	(0.09)	(0.17)	(-0.57)	(-0.86)
Above Group Mean (Top Tertile)	-3.055****	-1.724***	-2.790****	-0.642	-3.608****	-3.764****
	(-9.14)	(-2.68)	(-6.05)	(-0.86)	(-8.17)	(-3.88)
Below Group Mean (Bottom Tertile)	2.458****	3.556****	2.182****	3.513****	3.274****	3.997****
	(7.78)	(6.20)	(6.02)	(5.57)	(6.49)	(4.58)
Round (1–9)	-0.128***	-0.210**	-0.133**	-0.238**	-0.104*	-0.172
	(-2.99)	(-2.32)	(-2.49)	(-2.24)	(-1.72)	(-1.39)
Treatment(I)*Top Tertile		-2.288***		-3.641****		0.0543
		(-2.78)		(-3.71)		(0.04)
Treatment(II)*Top Tertile		-2.114**		-3.398***		0.0723
		(-2.29)		(-2.83)		(0.06)
Treatment(III)*Top Tertile		-0.855		-1.917		0.918
		(-0.92)		(-1.51)		(0.73)
Treatment(I)*Bottom Tertile		-1.492**		-2.104**		0.251
		(-1.96)		(-2.53)		(0.19)
Treatment(II)*Bottom Tertile		-2.105**		-2.296***		-2.338*
		(-2.46)		(-2.76)		(-1.82)
Treatment(III)*Bottom Tertile		-0.978		-0.888		-1.368
		(-0.97)		(-0.73)		(-0.89)
Treatment(I)*Round		0.136		0.158		0.116
		(1.26)		(1.21)		(0.67)
Treatment(II)*Round		0.00540		0.00606		-0.000109
		(0.04)		(0.04)		(-0.00)
Treatment(III)*Round		0.153		0.168		0.128
		(1.12)		(0.99)		(0.75)
Constant	0.403	0.0966	0.0910	-0.733	1.834	1.586
	(0.37)	(0.08)	(0.07)	(-0.56)	(0.81)	(0.63)
Observations	2680	2680	1672	1672	1008	1008

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Notes: OLS random effects estimation. z-statistics in parentheses. Robust standard errors, clustered at group level. Estimates for the pooled sample. Included demographic controls: age, sex, risk aversion, correct answer to control question, reading-time, CRT-score, and working memory score.

Summary Statistics

Table 15 contains summary statistics for the control variables used in all regressions above. As expected under random assignment, there are no significant differences between the BL and TP apart from working memory scores. This does not change when comparing slow and fast subjects in columns (4) - (6).

Table 15: Summary statistics by time pressure and compliance

	(1)	(2)	(3)	(4)	(5)	(6)
	BL	**TP**	BL vs. TP	**Slow**	**Fast**	Slow vs. Fast
	Mean (s.d.)	Mean (s.d.)	p-Value	Mean (s.d.)	Mean (s.d.)	p-Value
	N=172	N=176		N=196	N=152	
Age (Years)	22.71 (2.89)	22.83 (2.52)	0.38	22.65 (2.86)	22.93 (2.49)	0.13
Sex (1=Male)	0.46 (0.50)	0.49 (0.50)	0.58	0.46 (0.50)	0.49 (0.50)	0.53
Risk Aversion (1-11)	4.91 (2.43)	4.79 (2.37)	0.63	4.83 (2.44)	4.88 (2.34)	0.84
Unconfused (1=Yes)	0.52 (0.50)	0.58 (0.49)	0.24	0.53 (0.50)	0.58 (0.49)	0.44
Cognitive Reflection Score (0-3)	1.78 (1.10)	1.85 (1.12)	0.52	1.76 (1.12)	1.89 (1.10)	0.26
Readingtime Instructions (Log10 Sec.)	3.38 (0.38)	3.33 (0.42)	0.13	3.38 (0.41)	3.33 (0.39)	0.30
Working Memory Score (0-12)	4.64 (2.08)	5.45 (2.33)	< 0.01	4.56 (2.07)	5.68 (2.30)	< 0.001

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Notes: Individual characteristics by treatment assignment and treatment compliance. Pooled sample across order conditions. Fast subjects (Response times <= 7 seconds) and slow subjects (Response times > 7 seconds). P-Values in (3) and (6) are from M.W. ranks sum test for ordinal variables and from Chi^2 tests for binary variables.

Instructions

Experiment	Laboratory: Random seat assignment Personal code for anonymity Tasks implemented in z-Tree • General instructions *(page 45)* • Public Good Game HC *(page 45)* • Public Good Game CC *(page 50)* Payment according to personal code

General instructions
[SCREEN 1]
Dear participant, thank you for your participation. You will find general instructions concerning the procedure of the study on this screen:

- You will work through some computerized decision tasks and questionnaires. Please always follow the instructions on the screen in front of you. At the end of today's session, you are going to receive a monetary compensation. The funds for your compensation have been provided by the Ministry of Education and Research.
- As a compensation for your participation, you will receive €3. You will be able to earn additional money during the experiment. The exact amounts you will earn depend, on your own decisions during today's session.
- Every task will be explained to you. Please read the descriptions on the screen carefully.
- Of course all your decisions as well as your personal earnings from the experiment will be treated anonymously. The password you have created at the beginning of today's sessions serves to ensure this.

Introduction public good game
[SCREEN 2]

The first task is about to start. From now on please do not communicate with other participants in the room. In case you do so, we unfortunately will have to exclude you from the study. In this case you will not receive any compensation.

On the following screens you will find detailed instructions for the decision task. Please read them carefully. This ensures that you will know how to influence your earnings by your own decisions.

Instructions Public Good Game (HC)
[SCREEN 3]
Decision Task

Your main task in this study is to decide, how to divide 20 balls between two different bowls marked with **A** or **B**. You interact with **3 other participants** in this room. Thus including yourself, there are **4 participants** in a group. It will be impossible for you and all the other participants to observe who got matched with whom. Each of the other participants can also distribute the same number of balls (20) as yourself. You final payoff will depend on how you and the other participants distribute the balls between the two different bowls. The rules are identical for you and the other participants and all participants have received these instructions.

- **Bowl A:** Only you can fill bowl A. For each ball you put in your own bowl A, only you receive 20 Cent.

- **Bowl B:** You and the other 3 participants in your group can fill bowl B. The amount that you and all the other participants receive from bowl B depends on the total number of balls that are in bowl B. For each ball in bowl B you and each of the other 3 participants receive 10 Cent each.
- **The other 3 participants**: Each of the participants also receives 20 balls. For each ball that one of the other participants puts in his own bowl A, only he himself receives 20 Cent. For each ball that one of the other participants puts in bowl B, you, he and the other two participants receive 10 Cent each.

  So the payout rules are the same for all participants.
- **The final payoff:** Your final payoff depends on how you and the other participants fill the bowls. You will receive the payoff from **your bowl A**, as well as the payoff from the **joint bowl B**.

Procedure
[SCREEN 4]
Decision Task
Part I:
Overall, you will carry out the distribution task ten times.

First, you will take a decision only once. After stating your **first decision** you will receive new instructions that are only going to apply for the remaining nine decisions.

You will be matched anonymously with the same three participants in this room.

Part II:
After stating the first 10 decisions there will be a short questionnaire. After the questionnaire you will once again complete the distribution task for another 10 times.

For that purpose you will receive again new instructions. Please read these new instructions again carefully, as this can affect your earnings.

After the **first decision** round you will again receive additional instructions that are going to apply for the remaining 9 decisions.

Your final payoff:
At the end of this study, one of the 20 decisions is going to be selected at random. The probabilities for selecting a certain decision are the same (Like throwing a dice with the numers 1–20). Only this decision will be used to calculate your final earnings. So each decision is equally important for your final earnings.

End Instructions
[SCREEN 5]

You have completed all instructions and examples successfully.
You are now going to begin with the first 10 decisions.
(FOR TIME PRESSURE ONLY)
You have only a limited time budget available to enter your decision.

- Your time budget for the first 5 decisions is 7 seconds.
- For the second 5 decisions your time budget is 4 seconds.

For each decision in which you take longer than the time limit, 20 Cent will be deducted from your €3 show-up fee.
(FOR PEI ONLY)
You have only a limited time budget available to enter your decision.

- Your time budget for the first 5 decisions is 7 seconds.
- For the second 5 decisions your time budget is 4 seconds.

For the success of this experiment it is important that you make your decisions within this time limit
If you go over the time limit in a decision and this decision is selected for payment you will loose **your full show-up fee of €4.**
(FOR PRC ONLY)
You have only a limited time budget available to enter your decision.

- Your time budget for the first 5 decisions is 7 seconds.
- For the second 5 decisions your time budget is 4 seconds.

For the success of this experiment it is important that you make your decisions within this time limit
If you go over the time limit in a decision additional rules will apply that are explained to you on the next screen.
Additional rules (Shown on the following screen)

- The following rules only apply if you go over the time limit on a decision screen
- You will earn no money if this decision is selected for payment
- The other participants will still earn money from their own bowl A and bowl B
- From your 20 balls a random fraction will be placed in bowl B. The size of this fraction is predetermined and will not be communicated to you prior to your decision.
- You will not receive any payment from bowl A or B. The other participants, however, will receive a payment.

- It is therefore in your own interest to decide within the time limit.
- Your time budget for the first 5 decisions is 7 seconds.
- For the second 5 decisions your time budget is 4 seconds.

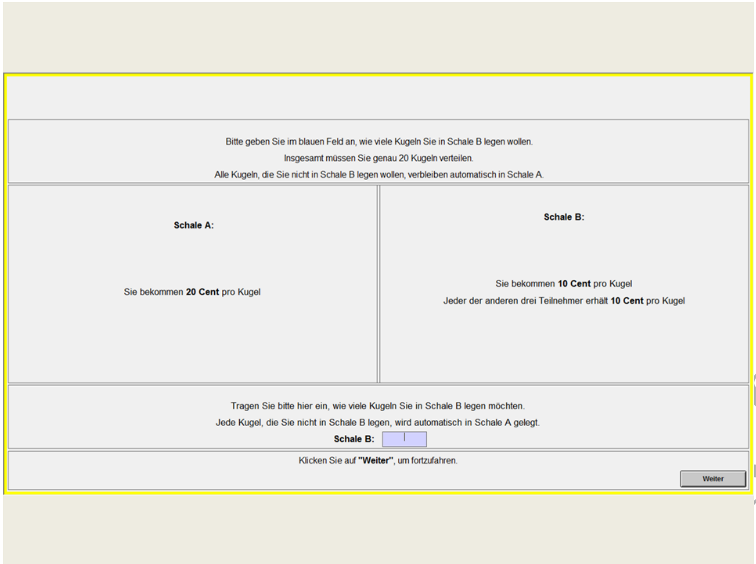


Figure 6: Screenshot PGG baseline/time pressure

Decision Screen
[SCREEN 6]
(FOR TREATED ONLY: Counter << +1 >>**)**
Please indicate in the blue field how many balls you want to put in bowl B. You have to distribute exactly 20 balls in total. All balls that you do not want to put in bowl B remain automatically in bowl A. You are free to choose any number of balls between 0 and 20.

- Bowl A: You receive 20 cents per ball.
- Bowl B: You receive 10 cents per ball. Each of the other 3 participants also receives 10 cents per ball.

$<< Entry : Contribuion(0-20) >>$

Additional Instructions
[SCREEN 7]
Additional rules for rounds 2-10

Additional information:

From now on you will be informed after each round how many balls the other participants have put into bowl B in total. The other participants that you interact with receive this information as well. The feedback screen will be left after a short time (10 Sec.) and the next round begins automatically.

Additional Decision Screens
Screens for decisions 2-10. Equivalent to screen 6.

Figure 7: Screenshot PGG Coputer Condition

Instructions Computer Condition (CC)
[SCREEN 8]

Description of payoffs equivalent.
Change of rules.
The other participants: As in the first 10 rounds, you will interact with three other players. However, these players **are not** other participants in this room. Instead these three players are controlled automatically by a computer program. Thus your interaction partners are no real human beings. Each of the three computer players has (like you) 20 balls that it divides up between bowl A and bowl B. The way the three computer players are going to divide up their balls between bowl A and bowl B has been determined prior to you first decision. Therefore, you cannot influence the computer players by your own choices. The contributions of the three computer players have been written on a poster here in this room that will be uncovered after your last decision at the end of the experiment. Thereby you can verify that the computers indeed act according to a preprogrammed contribution sequence.
While **you can earn actual money** from the balls in bowl A and B, the computer players naturally receive **no earnings** (as they are only a computer program)

Screen 10: Decision Screen Computer Condition
Please indicate in the blue field how many balls you want to put in bowl B. You have to distribute exactly 20 balls in total. All balls that you do not want to put in bowl B remain automatically in bowl A.

- **Bowl A:** You receive 20 cents for each ball.
- **Bowl B:** You receive 10 cents for each ball. Each of the other computer players “receives” 10 cents for each ball.

$<< Entry : Contribuion(0 - 20) >>$

Acknowledgments: The authors gratefully acknowledge financial support by the German Ministry for Education and Research under grant OIUV1012. They are furthermore thankful for helpful comments by seminar participants at the ESA Zurich, the HSC New York, the ZEW Mannheim, the University of Chicago, the University of Sterling, the IMEBESS Oxford and the SBRCC workshop in Kiel.