University of Heidelberg

Department of Economics

Discussion Paper Series | No. 626

Cooperation in Public Good Games. Calculated or Confused?

Timo Goeschl and Johannes Lohse

November 2016

Cooperation in Public Good Games. Calculated or Confused?

Timo Goeschl[a] Johannes Lohse[b]
University of Heidelberg University of Birmingham

November 24, 2016

Abstract

Recent experiments suggest that contribution decisions in a public goods game (PGG) are more likely to be cooperative if based on intuition rather than reflection. 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 subject's confusion status and a direct impact of time pressure on subjects' likelihood of being confused. Our reinvestigation on 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

1 Introduction

Economists' understanding of the drivers that underpin cooperation in the public goods game (PGG) has become increasingly sophisticated over the last twenty years (Ledyard, 1995; Chaudhuri, 2011; Vesterlund, 2014). Yet, despite the significant progress, it remains far from exhaustive and not without conflicting findings. As a result, researchers have adopted new conceptual approaches, some borrowed from psychology or evolutionary biology, as points of departure for novel experimental designs and new interpretations of existing experimental evidence. Among the most recent approaches, *cognitive style* has attracted significant attention as a possible cognitive driver of behavior in the PGG. Cognitive style refers to the type of mental processes that individuals draw on in order to come to a decision in a given choice situation. In economics, researchers have started to adopt so-called 'dual-system' models, which distinguish between intuition and deliberation as the two main cognitive styles (Loewenstein and O' Donoghue, 2007; Evans, 2008; Kahneman, 2011; Dreber et al., 2014).[1] In the context of social dilemmas, economists' recent interest in cognitive style as a driver of cooperation has at least three reasons. One is the potential of dual-system theories to provide additional building blocks for modeling cooperative behavior. A second is the possibility of designing decision environments that enhance cooperation once the links between cognitive style, cooperative behavior, and the situational determinants of cognitive style are better understood. A third reason is that differences in cognitive style could perhaps explain conflicting experimental findings that challenge existing outcome-based models of other-regarding preferences (Dreber et al., 2014).[2]

Cognitive factors have at least once before attracted the interest of economists studying subjects' behavior in the PGG. In his 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 *cognitive failure*: Confusion about core elements of the strategic situation was common among subjects in the PPG. Subsequent research has been able to confirm the important role of confusion for explaining positive contributions. It has also 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). In this paper, we connect the recent literature on cognitive style as an explanation of cooperative behavior in the PPG with the earlier literature on the role of cognitive failure among subjects playing that game form. These literatures not only share a common conceptual concern with cognitive factors in cooperation problems. More importantly, cognitive style and confusion are connected through the very methods commonly used to identify cognitive style. Because the cognitive processes involved in choice are not directly observable, researchers are beginning to collect supplementary non-choice data (Schotter, 2008; Schotter and Trevino, 2014), especially

[1] Apart from cooperation in the PGG, these models and the associated experiments have also been applied to other domains such as risky decisions (e.g., Kocher et al., 2013) or inter-temporal choice (e.g., Benjamin et al., 2013).

[2] In their paper, Dreber et al. (2014) describe a dual-self model of pro-social behavior that could reconcile some of these conflicting findings. For instance, standard social preference models cannot explain why a significant number of individuals chooses to avoid being asked to give, when there is a possibility to do so (Dana et al., 2007; Andreoni et al., 2011; DellaVigna et al., 2012). This behavior, however, would be consistent with a decision maker's deliberate choice to avoid the temptation of impulsive giving (Vesterlund, 2014; Dreber et al., 2014). Similarly, the difficulty to generalize from some lab results to other regarding behavior in the field (Levitt and List, 2007) could be due to the higher level of deliberation induced by the unfamiliar experimental environment.

on response times (Rubinstein, 2007; Spiliopoulos and Ortmann, 2014), or are putting subjects under time pressure, thereby exogenously influencing their cognitive style (Rand et al., 2012). Those and similar methods have been used separately or jointly in studies investigating the role of cognitive style in explaining outcomes of economic experiments in general and of social dilemma experiments in particular (Piovesan and Wengström, 2009; Hauge et al., 2009; Fiedler et al., 2013b; Ubeda, 2014; Grossman et al., 2014; Lohse et al., 2014; Corgnet et al., 2015; Achtziger et al., 2015). A highly cited example is a series of one-shot PGG experiments (Rand et al., 2012, 2014) in which the researchers found that participants deciding under time pressure contributed more, on average, than participants deciding in a time delay condition. Their interpretation of this evidence as proof of a causal link between intuition and cooperation and, more generally, as support for applying a dual-self model to cooperation problems generated significant attention in the literature (Gächter, 2012).

If decision speed exclusively correlates with cognitive style, but is unconnected with the presence of cognitive failure, evidence from time-pressure experiments can be safely interpreted without reference to economists' earlier insights on the role of confusion in PGG. But there are at least two channels through which cognitive failure could conceivably confound the inference from time pressure experiments. 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 when they come to the decision situation.[3] 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 instructions (Ferraro and Vossler, 2010).

The second 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 under time pressure. 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

[3]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."

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 behavior. We do so by emulating earlier studies that combine a linear PGG experiment with a time pressure treatment in order to reproduce the treatment effect of exogenously induced intuition on behavior. Reproducing the treatment effect and determining its direction is of interest in its own right: While several follow-up studies to Rand et al. (2012) affirm a positive relationship between intuition and cooperation based on correlational (Lotito et al., 2012; Nielsen et al., 2014) or causal evidence (Cone and Rand, 2014; Rand et al., 2015), a number of other studies find conflicting evidence. For example, Tinghög et al. (2013) and Verkoeijen and Bouwmeester (2014) fail to replicate the original findings. Using a different design, Duffy and Smith (2014) and Martinsson et al. (2014b) also find no evidence for intuitive cooperation in repeated public good games. A recent meta study (Rand, 2016) reports a significant positive average effect of intuition on cooperation (of 6.1 percentage points), but also shows that there are large differences across the single studies covered with many included studies not finding a significant difference in treatments. These conflicting findings call for more evidence on the presence and direction of the treatment effect. The second objective of our paper is 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 is 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 existing case for the role of cognitive style in explaining cooperation in PGG. But while the former would 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 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 must 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 robustness of the treatment effect to repeat play (Duffy and Smith, 2014; Martinsson et al., 2014b) is examined through a nine-round repetition run after the one-shot game.

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).[4] 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.

Our results are twofold. First, our reinvestigation on decision speed affirms that time pressure does affect behavior in the linear PGG with human partners (the HC). The direction of the effect, however, is the opposite of the findings by Rand et al. (2012), Cone and Rand (2014), and Rand et al. (2015) and supports the insignificant results of Tinghög et al. (2013) and Verkoeijen and Bouwmeester (2014): Comparing the BL and the TP conditions in the HC, we find that time pressure significantly increases the share of zero contributors and weakly decreases average contributions. This speaks against the conclusion that intuitive cognitive style favors cooperative behavior. The finding that faster subjects contribute significantly less to the public account is highly robust to potential selection effects.

Second, our evidence confirms that some, but not all concerns that the TP treatment effect could be confounded by confusion are justified. The nature of the confusion confound is a heterogeneous treatment effect only. In particular, contrary to the concerns raised in Recalde et al. (2014), we do not find that the method of using time pressure to induce a more intuitive cognitive style has the undesired effect of increasing cognitive failure, thereby confounding the true treatment effect. Time pressure itself does not affect the likelihood of cognitive failure, measured by positive contributions in the CC. On the other hand, there is a confound through a heterogeneous treatment effect, which 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 design to identify confused subjects, we find that time

[4] 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.

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. 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 contradict the earlier findings of the role of cognitive style on cooperation. These main findings on the effects of time pressure continue to hold when we move from a one-shot to a repeated setting, in which subjects can gain more experience with the task format. In sum, our results suggest that those contributions which can safely be interpreted as cooperative are rather the product of deliberation than intuition.
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* (NO) the first task was in the HC, while under *reverse order* (RO) 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 (HC-TP & CC-TP)

In the one-shot decision of the time pressure treatments, 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). In the later rounds of the repeated tasks (5-9) the limit was tightened to four seconds, to account for the possibility that subjects adapt to the time constraint. These 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. There are different approaches, how to deal with the possibility that subjects violate the time limit. One alternative would be a binding constraint, which shuts down the decision screen after reaching the time limit and automatically chooses a default contribution. We decided against a binding limit as this would complicate the game structure, by adding the option of strategic inaction for subjects in the TP condition. Instead, subjects could violate the time constraint. However, to reduce statistical problems associated with non-compliance (Tinghög et al., 2013), we introduced an incentive. For each violation of the time constraint subjects lost €0.20 of their show-up fee. For similar reasons, previous studies on cognitive load in economic games (including cooperation games) have also incentivised compliance (Cappelletti et al., 2011; Carpenter et al., 2013; Duffy and Smith, 2014). While no incentives were used in the original study by Rand et al. (2012) and hence their introduction might affect direct comparability of the treatment effects, we decided in favor of incentives given the high rate of non-compliance and the associated statistical problems in their original study.[5]

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 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 with sixteen or twelve subjects per session for a total of 348 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 (53%) to male (47%) participants.[6] 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 and all sessions were run by the same experimenter. 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 an incentivised comprehension question in which they had to state their payoff maximizing

[5] Also note that the cost 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 an effect on behavior in the PGG unlikely. For a more extensive discussion on the advantages and disadvantages of incentivising compliance in cognitive load tasks see Duffy and Smith (2014)

[6] Further summary statistics are contained in Table 9 of the Appendix.

strategy and a set of control questions.[7] 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 €9.51 (Min.:€4.80;Max.:€15.00), including a show-up fee of €3.

3 Results

We first discuss results from those one-shot PGGs, subjects encountered first in each condition (i.e., NO-HC and RO-CC). These outcomes are directly comparable to the evidence in Rand et al. (2012, 2014), Tinghög et al. (2013), and Verkoeijen and Bouwmeester (2014). We then proceed with the evidence from the reverse order condition and the repeated games to explore the role of experience and strategic interaction.

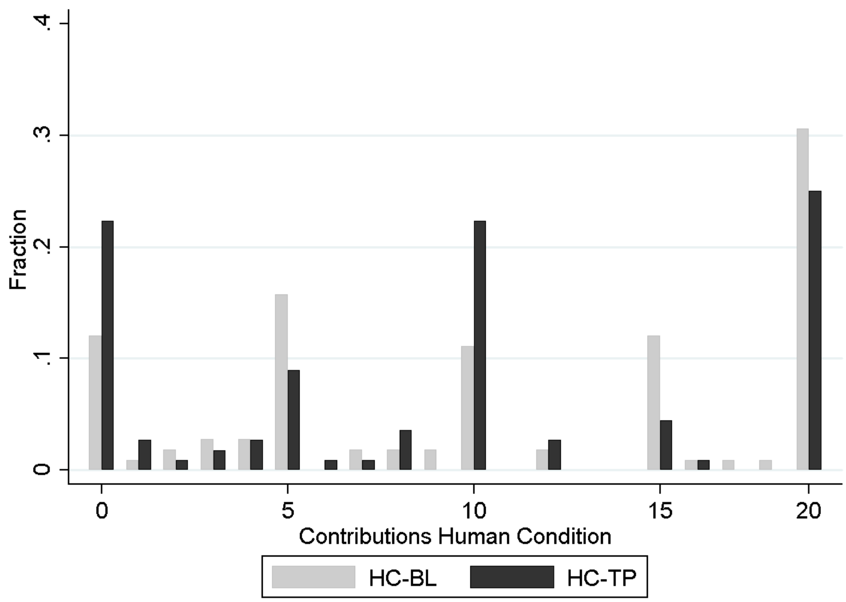
3.1 One-shot decisions

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*).[8] 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 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., 2014; 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 8 of the Appendix. Taken together, we therefore reject the hypothesis of intuitive cooperation and state the following result:

[7] 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). 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.

[8] 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.

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.[9] However, only 57.2 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, 2015). 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 (response times $\leqslant 7$ seconds) to slow subjects (response times > 7 seconds) 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,

[9] 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 7 of the Appendix.

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

Panel A: Second Stage (DV: Contributions)	(1) Contributions **HC**	(2) Contributions **HC**	(3) Free-Riding **HC**	(4) Free-Riding **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 (0-12)		0.230		-0.042
		(0.59)		(-1.04)
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 (0-12)		-0.027**		-0.027**
		(-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

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

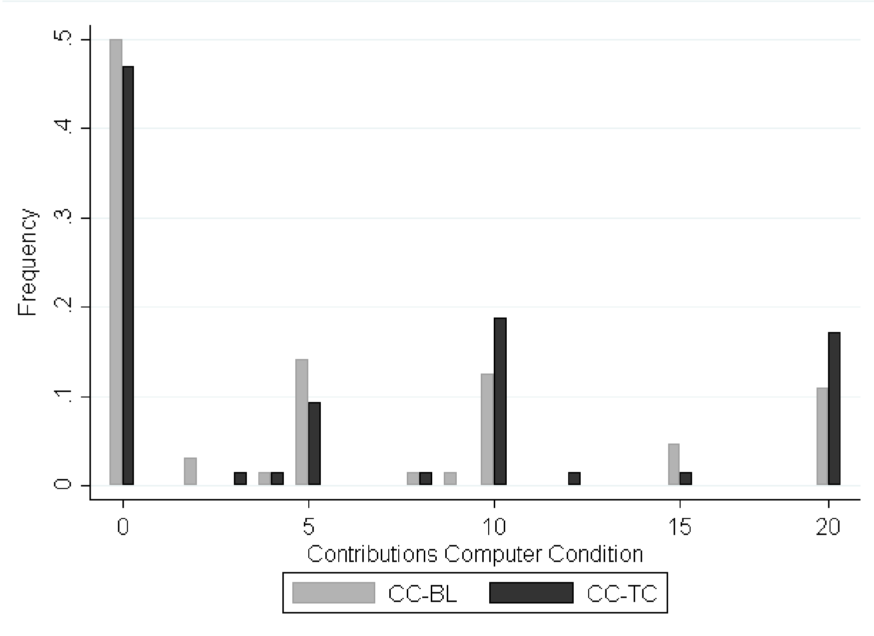
Notes: Specifications (1) and (2): Tobit-IV maximum likelihood estimation to account for censoring from below (0) and above (20). Specifications (3)-(4): Probit-IV maximum likelihood estimation. t-statistics in parentheses. Robust standard errors. Estimates for the pooled sample. Treatment effects are robust to using the following alternative specifications: using only observations from the normal order condition, using OLS instead of Tobit, using a dummy variable for fast decisions (response times either <= 5 or <= 7 seconds) instead of a continuous response time variable, using time pressure as the only instrument. The natural logarithm of response times is used to give less weight to outliers. Alternatively, excluding these outliers leads to equivalent results.

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.[10] 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 truly exogenous, while still being 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 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.[11] Both psychometric measures are not associated with the rate of free-riding (4).

[10] An overview over subject characteristics by treatment and compliance status is given in Table 9 of the Appendix.

[11] In a companion paper (Lohse, 2014) we explore and interpret this relationship more thoroughly using parts of the same dataset.

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 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=0.04 $p = 0.851$), nor do they contribute more on average (M.W. Rank Sum Test: z=-0.55, $p = 0.584$).[12] 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.

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

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.

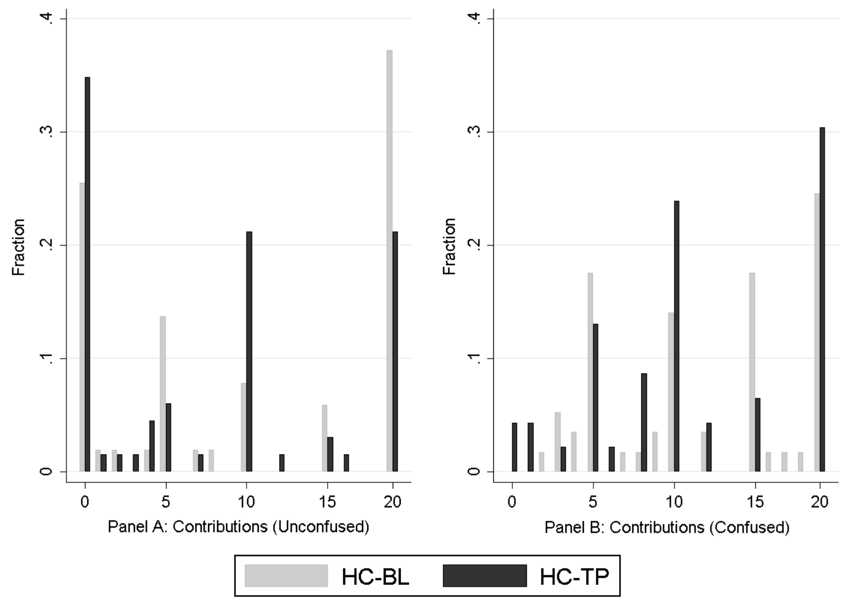
In our test for heterogeneous treatment effects we split the sample into "confused" and "unconfused" subjects. We do so by sorting a participant into the "confused" bin if we observe positive contributions in the one-shot game of the CC and into the "unconfused" bin if we observe zero contributions. 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 pressure on contributions between subjects in the "confused" and "unconfused" bins.[13] 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.[14] 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: 52.8%) 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. 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

[13] 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.

[14] 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.

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.

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.

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 (0-12)	0.704	0.954	0.085	0.197
	(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.

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 Repeated Decisions

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 repeated PGG, one in the human and one in the corresponding computer condition (compare Table 1). Prior to taking their first decision, 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. Based on the additional observations from the repeated games we 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.

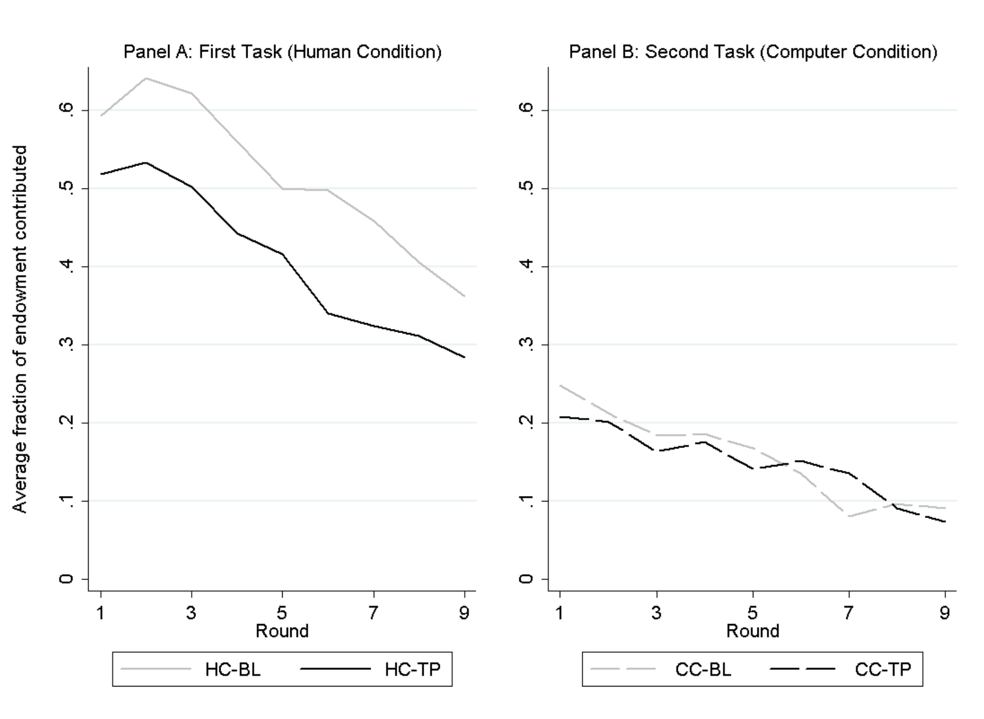
Table 4: Contributions averaged over nine rounds across treatment conditions

	(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)

Table 4 contains summary statistics on contribution rates averaged over all rounds. Between-subjects comparisons suggest that the treatment effects in the one-shot games are robust to repetition. The top row summarizes decisions from those repeated games which subjects encountered first under each condition. Consequently, subjects in these games have only been exposed to a limited amount of experience by deciding in the preceding one-shot game. We continue to 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 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 does neither affect 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, $p = 0.814$). Irrespective of task order, average contributions in the CC are significantly smaller than average contributions in the HC and overall confusion accounts for up to 40 percent of all tokens contributed in the human condition. This is slightly below the rates of confusion reported in Houser and Kurzban (2002) and Ferraro and Vossler (2010).

Figure 4 displays the evolution of average contributions (as a fraction of endowment) over time for 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. At this lower level of aggregation we continue to find no evidence for intuitive cooperation: 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, there is again no evidence, that time pressure affects 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

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

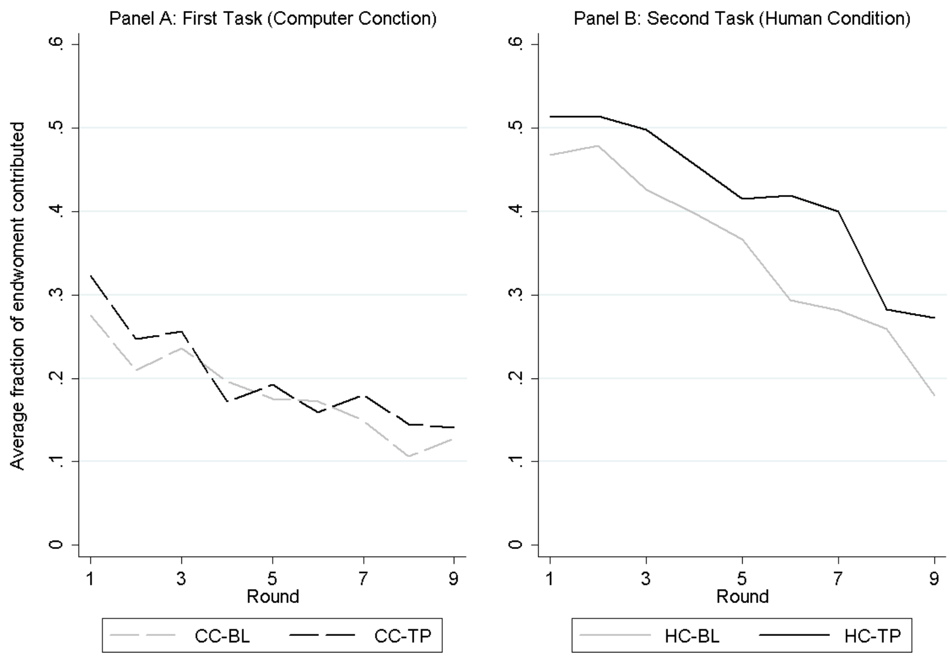


in the CC mostly represent a reduction of confusion, while declining contributions in the HC could additionally be due to "frustrated attempts" (Andreoni, 1995, p.892) at unreciprocated cooperation.

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 we continue to find 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 higher and the subsequent decline steeper than in the corresponding rounds from the CC conducted under normal order. Similarly, subjects deciding in the HC 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, in the HC there is no significant difference in average contributions between subjects in the baseline and subjects under timer 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 in Table 5 confirm the observations from Figures 4 and 5, when pooling data across both task orders. 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

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



three randomly assigned treatment conditions. 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.[15] In the CC there is 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

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

Table 5: 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	87	87	87
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. Estimates of treatment effects are robust to the following alternative specifications: Analyzing data at the group level, using Tobit models, clustering all standard errors at the individual level, estimating specifications from the HC and CC as seemingly unrelated regressions.

is weak evidence that subjects under time pressure converge slower in the early rounds, but display faster convergence in the last round. Taken together we state the following results:

Result 5: ***In the repeated games time pressure does not increase contributions. Instead, it significantly reduces contributions in the normal order condition and does not affect contributions in the reverse order condition.***

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

Finally, our data can shed some light on the question how deliberation and confusion interact to shape strategic behavior in a repeated setting. This question has not been addressed by previous time pressure experiments which have almost exclusively analyzed one-shot games. In a typical PGG about fifty percent of participants can be classified as conditional cooperators (Fischbacher et al., 2001; Chaudhuri, 2011): they increase (decrease) their own contributions to the public account from one round to another if (they believe that) the other group members also contribute more (less). The one-shot evidence in Rand et al. (2012, 2014) cannot be used to disentangle, whether time pressure affects conditional or unconditional cooperation in their setting (Gächter, 2012).[16] Furthermore, related 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. If time pressure has similar effects, we would expect to observe less conditional cooperation (defection) among treated subjects. 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.[17] Subjects contributing the same amount as the group average serve as the reference category.

Table 6 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

[16]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.

[17]While common in the literature, one problem with this empirical strategy could be that the behavior of other subjects cannot be seen as truly exogenous, as it might, in turn, depend on the past choices of the decision maker. Therefore, we devise a robustness check in which we only use data from the first two rounds. In these rounds the behavior of other subjects can be treated as exogenous, given that group composition is random. This robustness check arrives at similar conclusions.

Table 6: 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. Main results are robust to the following alternative specifications: estimating a Tobit model, using continuous instead of dummy variables to capture the behavior of group members.

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.[18] 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.[19] 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 (Kocher et al., 2012; Fiedler et al., 2013a; Tinghög et al., 2013; Verkoeijen and Bouwmeester, 2014; Martinsson et al., 2014a; Capraro and Cococcioni, 2016), we reinvestigate the initial claim that a more intuitive cognitive style causes subjects 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 (Rand et al., 2014).

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

[18] 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).

[19] 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.

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 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. 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 in line with findings in Strømland et al. (2016), who also identify confusion as key moderator. This heterogeneous treatment effect could be a 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 contributions 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, J. M. Rao, and H. Trachtman. Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *NBER Working Paper No. 17648*, 2011.

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.

D. J. Benjamin, S. A. Brown, and J. M. Shapiro. Who is behavioral? Cognitive ability and anomalous preferences. *Journal of the European Economic Association*, 11(6):1231–1255, 2013.

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

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. *NHH Dept. of Economics Discussion Paper*, (09), 2014.

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. *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.

J. Dana, R. A. Weber, and J. X. Kuang. Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1):67–80, 2007.

S. DellaVigna, J. A. List, and U. Malmendier. Testing for altruism and social pressure in charitable giving. *The Quarterly Journal of Economics*, 127(1):1–56, 2012.

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.

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.

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.

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. Are social preferences skin deep? Dictators under cognitive load. *GUPEA Working Papers in Economics 371*, 2009.

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, P. Martinsson, K. O. R. Myrseth, and C. E. Wollbrant. Strong, bold, and kind: Self-control and cooperation in social dilemmas. *Experimental Economics*, pages 1–26, 2012.

M. G. Kocher, J. Pahlke, and S. T. Trautmann. Tempus fugit: Time pressure in risky decisions. *Management Science*, 59(10):2380–2391, 2013.

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.

S. D. Levitt and J. A. List. What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives*, 21(2):153–174, 2007.

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. *University of Heidelberg Department of Economics Discussion Paper Series*, 578, 2014.

J. Lohse, T. Goeschl, and J. Diederich. Giving is a question of time: Response times and contributions to a real world public good. *University of Heidelberg Department of Economics Discussion Paper Series*, (566), 2014.

G. Lotito, M. Migheli, and G. Ortona. Is cooperation instinctive? Evidence from the response times in a public goods game. *Journal of Bioeconomics*, 15(2):1–11, 2012.

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

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

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. Myrseth and C. Wollbrant. Intuitive cooperation refuted: Commentary on Rand et al. (2012) and Rand et al. (2014). *GUEPA Working Papers in Economics*, 617, 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. *Working Paper (Available at SSRN 2401325)*, 2014.

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.

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.

L. Vesterlund. Voluntary giving to public goods: moving beyond the linear VCM. In J. Kagel and A. Roth, editors, *Handbook of experimental economics*, number 2. Priceton University Press, Priceton, 2014.

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

Appendix

Response Time Distribution

Table 7 summarizes the distribution of response times in the one-shot public good game across the four HC.

Table 7: 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.

Regression results: treatment effects one-shot public good game

In tobit regression models (1) and (2) of Table 8 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.

Table 8: Effects of treatment assignment

	(1) Contributions	(2) Contributions	(3) Free-Riding	(4) Free-Riding
Treatment(I): Normal Order + Time Pressure (1=Yes)	-3.256*	-3.270*	0.412**	0.430**
	(-1.80)	(-1.83)	(2.01)	(2.04)
Treatment(II): Reverse Order (1=Yes)	-3.384	-3.298	0.343	0.406*
	(-1.61)	(-1.62)	(1.44)	(1.67)
Treatment(III): Reverse Order + Time Pressure (1=Yes)	-4.820**	-5.571***	0.594***	0.653***
	(-2.22)	(-2.64)	(2.60)	(2.73)
Age (Years)		-0.228		-0.009
		(-0.86)		(-0.27)
Sex (1=Male)		3.390**		-0.132
		(2.33)		(-0.80)
Risk Aversion (1-11)		0.244		-0.008
		(0.82)		(-0.24)
Unconfused (1=Yes)		-1.867		0.209
		(-1.26)		(1.25)
Cognitive Reflection Score (0-3)		1.840***		0.055
		(2.76)		(0.73)
Readingtime Instructions (Log10 Sec.)		4.171**		-0.165
		(2.49)		(-0.81)
Working Memory Score (0-12)		0.102		-0.028
		(0.27)		(-0.71)
Constant	12.82****	-1.407	-1.173****	-0.443
	(10.01)	(-0.16)	(-7.50)	(-0.41)
Observations	348	335	348	335

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

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. Results are robust to using the following alternative specifications: using only observations from the normal order condition.

Summary Statistics

Table 9 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 9: 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 35)* • Public Good Game HC *(page 35)* • Public Good Game CC *(page 38)* 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.

Figure 6: Screenshot PGG baseline/time pressure

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 round in which you take longer than the time limit, 20 Cent will be deducted from your €3 show-up fee.

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.

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.

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

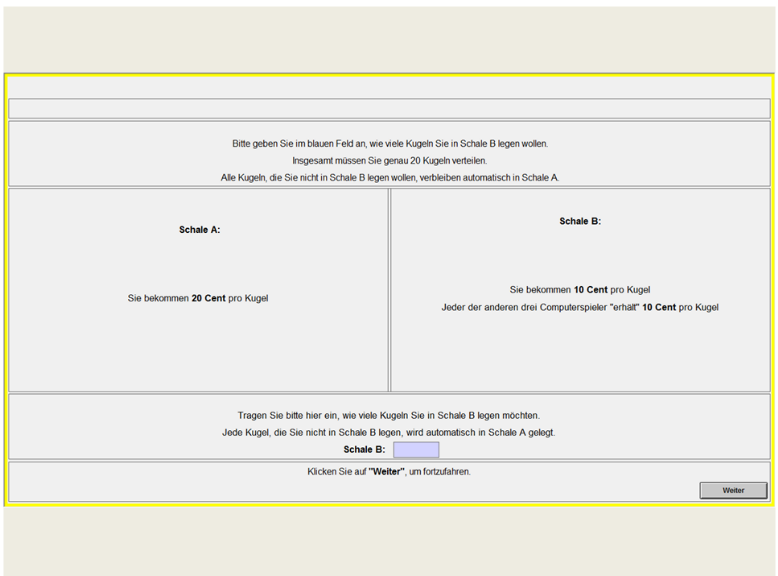


Figure 7: Screenshot PGG Coputer Condition

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.