# Cognitive Uncertainty, GPT, and Contribution in Public Goods Game*

Te Bao Jiaoying Pei

25th April 2024

**Abstract**

This paper establishes a connection between cognitive noise (Enke and Graeber, 2023) and the level of contribution in the public goods game. We argue that cognitive noise complements, rather than replaces, taste-based social preference to explain the contribution decision. Both correlational and causal data supports the notion that cognitive uncertainty is positively correlated with contribution in the public goods game at the aggregate level, or cognitive uncertainty led people to behave as if they are more cooperative. And the result is robust when removing strategic uncertainty. However, there is heterogeneity, where cognitive noise is negatively correlated with the contribution level of some participants at an economically significant extent. These findings suggest the significance of only considering contribution decisions that exceed a certain cognitive certainty threshold in a public goods game if they are to be taken at face value. Further, our experimental results also demonstrate that a cooperative advice from the Generative Pre-trained Transformer (hereafter referred to as "GPT") reduces cognitive uncertainty for all participants assist individual in either gaining a better understanding of their true social preference, or translating their true social preferences into contribution actions that maximize their utility as the game repeats. The impact of the advice, however, does not seem to depend on whether or not the participants are informed the advice was made by GPT.

**Keywords:** Cognitive Noise Experiments, Public Goods Experiments, GPT
**JEL:** C92, D91, H41, G0

*We are grateful to Benjamin Enke, Lu Dong, Peiran Jiao, Arno Riedl, Theo Offerman, Ning Sun, Erik Snowberg, Jan Tuinstra, Marie Claire Villeval, Florian Wagner, Bin Xu, Yilong Xu, Maoliang Ye, Xu Zhang, Xinxin Zhu and seminar audiences at East China Normal University, Fudan University, Southern University of Technology, Sun Yat-Sen University, Tinbergen Institute, Zhejiang Gongshang University for very valuable comments and discussions. This study is approved by IRB of Nanyang Technological University (IRB-2022-834). Financial support NTU CoHASS Seed Grant is gratefully acknowledged. Te Bao: School of Social Sciences, Nanyang Technological University, Singapore, Email: baote@ntu.edu.sg; Jiaoying Pei: School of Social Sciences, Nanyang Environment & Water Research Institute, Environmental Process Modelling Centre, Interdisciplinary Graduate Programme, Nanyang Technological University, Singapore, Email: peij0003@e.ntu.edu.sg.

# 1 Introduction

The public goods game (Isaac et al., 1984; Isaac et al., 1994; 1998) has been widely used to study cooperative behavior in groups when players have incentives to free ride. In the public goods game, individuals can contribute money to a public account that produces benefits shared by all group members, including themselves. When there is no opportunity for punishment or reciprocity concerns, experimental results show that the contribution rate decreases from 30% to 40% to 10% to 20% as the game repeats. Previous studies usually use social preferences, e.g., altruism, inequality aversion, indirect reciprocity to explain why people cooperate even when zero contribution is their monetary payoff maximization action (Fehr and Fischbacher, 2003).

In this paper, we hypothesize that the cognitive uncertainty[1] that is recently defined by Enke and Graeber (2023) could *complement* the taste-based social preference to together explain the contribution decision in public goods game.

Our approach is similar as how Enke et al. (2023) recently apply cognitive noise to intertemporal choice and explains the various empirical regularities in intertemporal choice[2]. In their theoretical and experimental study, the cognitive noisiness induces a "compression effect" that make people behave as if they treat different time delays to some degree alike, which leads to an inelasticity of decisions but not replaced the taste-based present bias. More specifically, cognitive uncertainty represents people's internal uncertainty about their utility-maximizing action, assuming that people are aware that they may make mistakes in their decisions, or in other words, how "good" their decision represents their true preference[3]. By constructing a Bayesian updating model and conducting a series of experiments that measure and manipulate cognitive uncertainty, they prove their hypothesis that people make decisions using a weighted function between their true time preference and a cognitive default or an ignorance prior. When subjects are more cognitively uncertain about a decision, more weight is assigned to the cognitive default (e.g., the middle of the response scale),

[1] The cognitive uncertainty here is similar to the noise proposed in Burton-Chellew and West (2013), which suggest that in the context of a public goods game, it could result from errors, boredom, learning, exploration, fluctuating preferences, or evolutionary constraints. The cognitive uncertainty is, however, different from confusion in Burton-Chellew et al. (2016), where they argue that contribution in public goods game is from players' confusion that is defined as their failure to understand the nature of the game. In other words, confusion there is the failure to understand that contributing zero endowment is the payoff-maximization action that is unconditional on their social preference. Recently, Wang et al. (2024) replicate the experiment in Burton-Chellew et al. (2016) with the only difference to minimize confusion by providing participants with increased training, i.e., by telling participants the correct answers to the 10 control questions. Their result refuted that in Burton-Chellew et al. (2016). Rather, their result suggest that the cooperation is not pure artifact of confusion and supports the notion that choices in public goods game reveal motivations such as adherence to social norm.

[2] The empirical regularities they explain using cognitive noise include present bias, hyperbolic discounting, and why choices frequently violate transitivity.

[3] Note that the internal uncertainty here distinguishes them from external uncertainty that, in contrast, roots in the environment, e.g., the risk of not receiving the payment. It also differentiates from the external outcome uncertainty — the uncertainty about whether or not a decision will lead to a particular outcome. It further differentiates from the impact uncertainty — an uncertainty about how badly others' well-being will be impacted by the negative outcome. Using an experiment, Kappes et al. (2018) find when the external outcome is more uncertain, people are less prosocial; but when the impact of subjects' action to opponent is more uncertain, people become more prosocial.

resulting in observed decisions that deviate more significantly from their true time preference.

This paper contains 3 self-contained parts to link the concept of cognitive noise with the level of contribution in public goods game. Our main idea is that cognitive noisiness leads to a distortion on the taste-based social preference measured by the public goods, but not replace it. Despite the many literatures investigating whether cooperation from public goods game stems from either confusion or prosocial behavior, we believe our paper is the first that accepts the premise that people can both possess prosocial preference, and be confused about the utility-maximizing decision that best represents their true social preference.

In Part 1 of this paper, we hypothesize that subjects will become increasingly cognitively certain about their decisions as the game repeats, due to experience and learning. Cognitive uncertainty can arise from either unawareness of one's true social preference, or uncertainty about the optimal action to maximize utility given their social preference in a public goods game setting. Consequently, subjects' contribution levels in the later rounds of the game will deviate more from a cognitive default that is at half of the response scale, as less weight is assigned to the cognitive default. To test this hypothesis, we simply observe the correlational change in cognitive uncertainty as the game repeats. Then, we examine whether the contribution rates stay closer to half of the endowment at the beginning of the game while deviating more from the half-of-endowment-level at the later rounds of the game.

The experimental result in Part 1 further allows us to ask how cognitive uncertainty would bias one's level of contribution, and in turn, bias the apparent social preference. In Part 2, and first through a correlational study, we first observe whether cognitive noise makes people appear to be more cooperative (less cooperative) on the aggregate level. This is done by comparing whether individuals who are more cognitively uncertain contribute more (less) on average than those who are cognitively certain about their decisions. Then, we determine the proportion of subjects where cognitive noise makes them appear to be more cooperative (less cooperative), but in reality, they are less cooperative (more cooperative), as they contribute less (more) when cognitive noise decreases. In Appendix E, and as a causal study, we implement additional treatment arms in which we complicate public goods game decision tasks by displaying the efficiency factor into a mathematical equation to increase confusion (Enke et al., 2023; Andreoni, 1995). To prevent subjects taking longer time to complete tasks which offset the effect of complexity, we also implement a time limit of 25 seconds per contribution decision. We then observe whether the result from correlational study holds in this causal study.

We conduct the cognitive noise experiments (Enke and Graeber, 2023) in the context of the without-punishment public goods game of Fehr and Gächter (2000). There are a total of 10 periods of the game in each part. In each period, each participant is endowed with 20 points, and they can choose to contribute to a project with a multiplier of 1.6. The benefits from this project, 1.6 unit for every unit of contribution, will be equally divided among the 4 group members. To rule out reciprocity concerns, participants are paired with different individuals in each period. Furthermore, in order to mitigate the influence of inequality aversion, participants are only able to view their own payoff, and a time limit of 10 seconds is enforced in each period to prevent them from calculating the payoffs of others. To measure cognitive uncertainty, we include a cognitive uncertainty question for each decision.

Our results demonstrate that only a cooperative advice from Generative Pre-trained Transformer (hereafter referred to as "GPT") can assist individual in either gaining a better understanding of their true preference or translating their true preferences into contribution actions that maximize their utility as the game repeats. The impact of the advice, however, does not seem to depend on whether or not the participants are informed the advice was made by GPT, implying an absence of GPT premium. By contrast, we fail to find a decreasing cognitive uncertainty as game repeats in the context when subjects do not receive any advice before making decisions.

Further, we argue that cognitive noise complements, rather than replaces, taste-based social preference to explain the contribution decision. Using both correlational and causal study, we find that cognitive noise distorts the contribution decision that primarily reflects subjects' true social preference. Across all treatments, we observe that subjects assign more weight to the cognitive default (50% contribution of their endowment, implying a middle bias) when they are less cognitively certain about their decisions, and therefore support the cognitive uncertainty hypothesis proposed by Enke and Graeber (2023).

At the aggregate level, cognitive uncertainty leads individuals to contribute more to the public account, which means that they may be less cooperative than what their decisions imply. We also find heterogeneity among the participants in terms of the direction in which cognitive noise biases their apparent social preference. While cognitive noise is positively correlated with the contribution level of the majority of subjects, there is still a minority of subjects whose level of contribution is negatively correlated with cognitive noise to the extent that is economically significant. In other words, there are a minority of subjects in our experiment are in fact more cooperative, while cognitive noise makes them appear to be less cooperative. These results suggest that when researchers are interested in determining subjects' true social preference, they should consider their contribution decisions only when the corresponding cognitive uncertainty is below an acceptable threshold.

As a robustness check and to address concerns that strategic uncertainty (Enke and Graeber, 1988; Gangadharan and Nemes, 2009) may be misinterpreted as part of the cognitive uncertainty being measured, we design Experiment Robustness that disentangle the two types of uncertainty in Appendix F. i) Strategic uncertainty, where they are unclear about their opponent's types and, consequently, their opponents' contributions, and ii) Cognitive uncertainty, where participants are unsure about their prosocial preference, or the utility-maximizing action in the public goods game that represents their true preferences. We randomly assign a subset of subjects into the Full

Information Treatment, so that they will always know the contribution of all other group members before making their own decision. Meanwhile, they will always be paired with three other subjects who do are not in the Full Information Treatment, and the other subjects do not know that they are playing with a subject who can access to their contribution decision.

We find that when ruling out strategic uncertainty, the reported cognitive uncertainty is different from zero at a statistically significant level. And the observation where subjects contribute more when they are more cognitively uncertain still holds in this scenario.

Recent research shows that GPT outperforms humans in rationality in decision-making tasks concerning risk, time, social, and food (Chen et al., 2023). With the increasing popularity of online platforms for conducting economic experiments to elicit preferences, along with the growing prevalence of GPT usage, concerns arise about the unbiasedness of the apparent preferences elicited in online economic experiments. Specifically, in the absence of supervision, subjects may delegate the decision to GPT instead of exerting cognitive effort and providing an answer that reflects their true preferences. Fortunately, the recent experimental evidence from Bai et al. (2023) find that individuals prefer employing AI to empower their judgements rather than entirely delegating the decisions to AI.

In Study 3, we investigate the following questions: How does subjects' contribution decision change after being exposed to decisions from GPT, and how long does the effect sustain? Does an exposure to GPT reduce subjects' cognitive uncertainty regarding their contribution decisions in a public goods game? Furthermore, does the extent where participants trust and adopt recommendations depend on whether or not they are informed that the recommendations come from GPT?

To answer these questions, we design two additional treatment group: Group Adviser and Group GPT. Note that we adopt a between + within treatment design for all the treatments in the main study of this paper[4]. All subjects are placed into groups of four and play for 20 rounds, with a restart prior to Rounds 11. In Group Baseline, the two rounds of the game are exactly the same. In contrast, participants in Group Adviser and Group GPT are provided with the decisions from GPT before the new 10 rounds of the game. We explicitly inform them that these decisions are from GPT in the Group GPT, and subjects are provided with a link directing them to the Wikipedia page of GPT-3.5. However, we only mention that the decisions are collected from an adviser in the Group Adviser. This allows us to observe how much subjects trust GPT compared to an anonymous adviser.

Our experimental results support the hypothesis that subjects' responses may be biased towards GPT's decisions in the absence of supervision during online experiments. However, the impact

[4]This means that whatever the treatment group subjects are allocated in, they will always play the standard public goods game for the first 10 rounds. And the treatment is only imposed after the surprise restart of the game at round 11. See Table B.2 for details.

form GPT is relatively short run and un-sustained for 10 rounds of the game. This finding is consistent with the experimental evidence in Bai et al. (2023): Even if equipped with a punishment system based on scoring the contribution behavior from AI or even real participants in the group, high contributions will start to decay before the end of each of the 10 rounds of the game (See Figure 1 in Bai et al., 2023 for details). But in the short run, whether they specifically bias their decisions towards the explanation provided by GPT cannot be simply captured by changes in cognitive uncertainty, at least in the context where GPT advises a cooperative contribution decision. These findings suggest that future studies aiming to explore subjects' preferences through online experiments will need to conduct additional supervised laboratory experiments for robustness. Meanwhile, we did not observe any significant and sustained difference between Group Adviser and Group GPT in both cognitive uncertainty and contribution decision, indicating the absence of a GPT premium. Our result provides a useful perspective on how AI-human interaction influences human decision on the social context.

Our study is closely related to Andreoni (1995). They designed a lab experiment to separate the hypothesis that the contribution in the public goods game is due to kindness, or simply the result of errors or confusions. They calculated the percentage of contributions that fall into either of these two categories (in the way lists on pp. 895 paragraph 1). As half of the cooperation come from subjects who understand free-riding but still choose to cooperate, implying that they belong to kindness category, they argue that there exists non trivial amount of people who possess a cooperative preference. Further, their experimental evidence concludes that the declining contribution may due to frustrated attempts at kindness instead of learning.

Our paper is aligned with the literature that quantitatively captures confusion by employing the cognitive uncertainty concept proposed by Enke and Graeber (2023). In other words, rather than categorizing contribution behavior into pure kindness or pure confusion like they do, we accept the premise that subjects can exhibit both qualities at the same time. And more importantly, we view each contribution decision as a weighted function of the two qualities and focus on how confusion can influence their observed social preferences, i.e., contribution decision. Our results, although derived from this different approach, support the findings of Andreoni (1995). Similar to their conclusion, we find that cognitive uncertainty affects the contribution decision in the public goods game, but it cannot solely explain it. Specifically, there exists a non-trivial number of subjects whose confusion or cognitive uncertainty makes them appear to be less cooperative or less kind, suggesting that they are inherently more cooperative than what is observed.

We contribute to the recent literature on cognitive noise experiments, which explore the concept of cognitive uncertainty and its relationship with various empirical regularities, including risk, ambiguity, belief updating, survey forecasts of economic variables (Enke and Graeber, 2023), intertemporal choice (Enke et al., 2023), and overconfidence (Amelio, 2022).

Further, we contribute to the literature on public goods experiments (e.g., Isaac et al., 1984;

1994; Fehr and Gächter, 2000; Fehr and Fischbacher, 2002; Fischbacher and Gächter, 2006; Burton-Chellew and West, 2013; Burton-Chellew et al., 2016; Bao et al., 2022; Wang et al., 2024) and the study of the restart effect in the game (e.g., Croson, 1996; Burton-Chellew, 2022). Specifically, we link the noise behind the non-zero contribution in the public goods game hypothesized by Burton-Chellew and West (2013) with the cognitive noise defined by Enke and Graeber (2023) and test it through economic experiment. We find that cognitive noise introduces bias in contribution decisions. While cognitive noise amplifies contributions at the aggregate level, there is heterogeneity within the population regarding the direction in which it influences contributions.

Relatedly, our study also provides a mechanism for explaining why an increase in time pressure (e.g., Rand et al., 2012; 2014; Rand and Kraft-Todd, 2014; Cone and Rand, 2014) or cognitive load (e.g., Schulz et al., 2014; Cornelissen et al., 2011; Roch et al., 2000) could increase cooperation. Adapting the cognitive noise model from Enke and Graeber (2023), our experimental results show that manipulating complexity together with time pressure (even when it is not binding) would increases subjects' cognitive load. This, in turn, causes subjects to assign less weight to their true social preferences and more weight to a cognitive default (i.e., a middle bias) when making contribution decisions. Since most subjects would contribute less-than-half-of-their-endowment when they are cognitively certain, cognitive uncertainty makes the subjects appear more cooperative.

Another contribution of our study is the additional tracking of the dynamics of cognitive uncertainty in the public goods game. We find that the dynamics of cognitive uncertainty depend on whether subjects receive cooperative advice before making their decisions. Specifically, in the setting of the classic public goods game, there is no clear trend in cognitive uncertainty as the game repeats. However, when subjects were provided with cooperative advice before making their decisions, regardless of whether the advice came from an anonymous adviser or GPT, they showed a decreasing cognitive uncertainty in their decisions as the game repeated. Generally speaking, our experimental results demonstrate that a cooperative advice can assist individual in either gaining a better understanding of their true social preference or translating their true social preferences into contribution actions that maximize their utility as the game repeats.

Finally, we are related to the recent literature exploring rationality (Chen et al., 2023), social preferences (Guo, 2023), and other capabilities (e.g., Brown et al., 2020; Chen et al., 2021; Lin et al., 2020; Drori et al., 2022) of large language models (LLM) and GPT. We find that GPT-3.5 usually recommends a fairly cooperative level of contribution when instructed to act as a human decision maker in a public goods game.

The paper is organized as follows: Section 2 outlines the experimental design. Section 3 presents the results. Finally, Section 4 concludes the paper.

# 2 Experimental Design

## 2.1 Testable Hypotheses

This paper aims to explore the relationship between cognitive noise and contribution decisions in the public goods game.

Specifically, we examine whether the observed decline in contributions, which is often reported in the literature in the without-punishment public goods game, can be attributed to a decrease in cognitive uncertainty during the decision-making process.

First, we investigate whether participants become more certain about their decisions as the game repeats, where cognitive uncertainty arising from either being unaware of their true social preference or from uncertainty regarding the optimal action to maximize utility based on their social preference. We formulate our first hypothesis as follows:

**Hypothesis 1:** *Cognitive uncertainty decreases over time. As the game repeats, participants become more certain about their contribution decisions.*

According to the cognitive noise hypothesis proposed by Enke and Graeber (2023), when individuals experience higher cognitive uncertainty in their decision-making, they tend to assign more weight to the cognitive default (e.g., the middle of the response scale). Consequently, this leads to observed decisions that deviate more significantly from their true preferences. Combining this with Hypothesis 1, we propose our second hypothesis:

**Hypothesis 2:** *Higher measured cognitive uncertainty is associated with decisions closer to a 50% contribution of the endowment. This implies that as the game repeats, the absolute difference between the actual contribution and the 50% contribution of the endowment becomes larger.*

Considering the decreasing trend typically observed in the literature and our hypothesis of an increasing cognitive uncertainty as game repeats, we also hypothesize a negative correlation between cognitive uncertainty and contribution decisions.

**Hypothesis 3:** *Cognitive noise is positively correlated with contribution. In other words, cognitive uncertainty leads individuals to behave as if they are more cooperative.*

Lastly, we hypothesize that receiving a cooperative recommendation from an advisor or the GPT would alter participants' cognitive uncertainty about their decision, thereby influencing their ultimate decision. We further hypothesize that the degree to which participants trust and adopt advice from a general advisor and the GPT will differ.

**Hypothesis 4:** *Participants' cognitive uncertainty changes when they receive a cooperative recommendation, and their decision also leans towards the cooperative recommendation. Furthermore, participants may trust and adopt recommendations to different extents depending on whether or not* *they are informed that it comes from GPT.*

### 2.2 Participation Pool, Logistics, Group, and Treatment

The experiments were conducted online with undergraduate students recruited from Nanyang Technological University (NTU), and programmed using oTree (Chen et al., 2016). A total of 140 NTU students from all majors participated in our experiments on June 22, 2023. At the beginning of each session, each subject received their private experimental link through a direct private chat with the experimenter on Microsoft Teams. All participants do not know the size of the session or the identity of other participants.

The experimental procedure is as follows, with the detailed instructions can be found in Appendix C. Once participants had finished reading the instructions, they were instructed to complete four control questions on the public goods game from Fehr and Gächter (2000), as well as one control question on cognitive uncertainty from Enke and Graeber (2023). Once they answered all questions correctly, subjects were then randomly paired into groups of four to play the public goods game and indicate their level of certainty regarding their contribution decision after each round. Specifically, after making the contribution decision for each round, the subsequent screen reminded them of their previous decision and elicited cognitive uncertainty. Participants answered the following question by selecting a radio button between 0% and 100% in steps of 10%.

> *Your decision on the previous screen indicates that you would like to contribute $x$ points to the group project. How certain are you that you would actually want to contribute somewhere between $x - 1$ points and $x + 1$ points to the project?*

We isolate reciprocity by inform subjects that there is reshuffling the composition of the group in each round. To also eliminate the effect of inequality aversion, subjects were only able to see their own payoff, and there was a time limit of 10 seconds in each period for viewing the payoff to prevent them from calculating others' payoff.

All subjects are placed into groups of four and play for 20 rounds, with a restart prior to Rounds 11. They were informed that the instructions for the next 10 rounds were exactly the same as the previous 10 rounds, which were shown to them as a reminder. The only difference was that subjects in Group GPT were prompted with a recommendation by GPT-3.5, informing them that GPT usually contributes 13 out of 20 of the endowment[5], based on the same instruction provided to them. Subjects in the Group Adviser were given the same recommendation, with the distinction that they were told the recommendation came from an adviser.

[5]The decision from GPT was generated using the GPT API, with the instructions largely adapted from (Chen et al., 2023) where they investigated the rationality of GPT. On average, GPT contributed 13 out of 20 points in the 150 decisions it made. The detailed instructions to GPT and the summary statistics of GPT's decisions can be found in Appendix D.

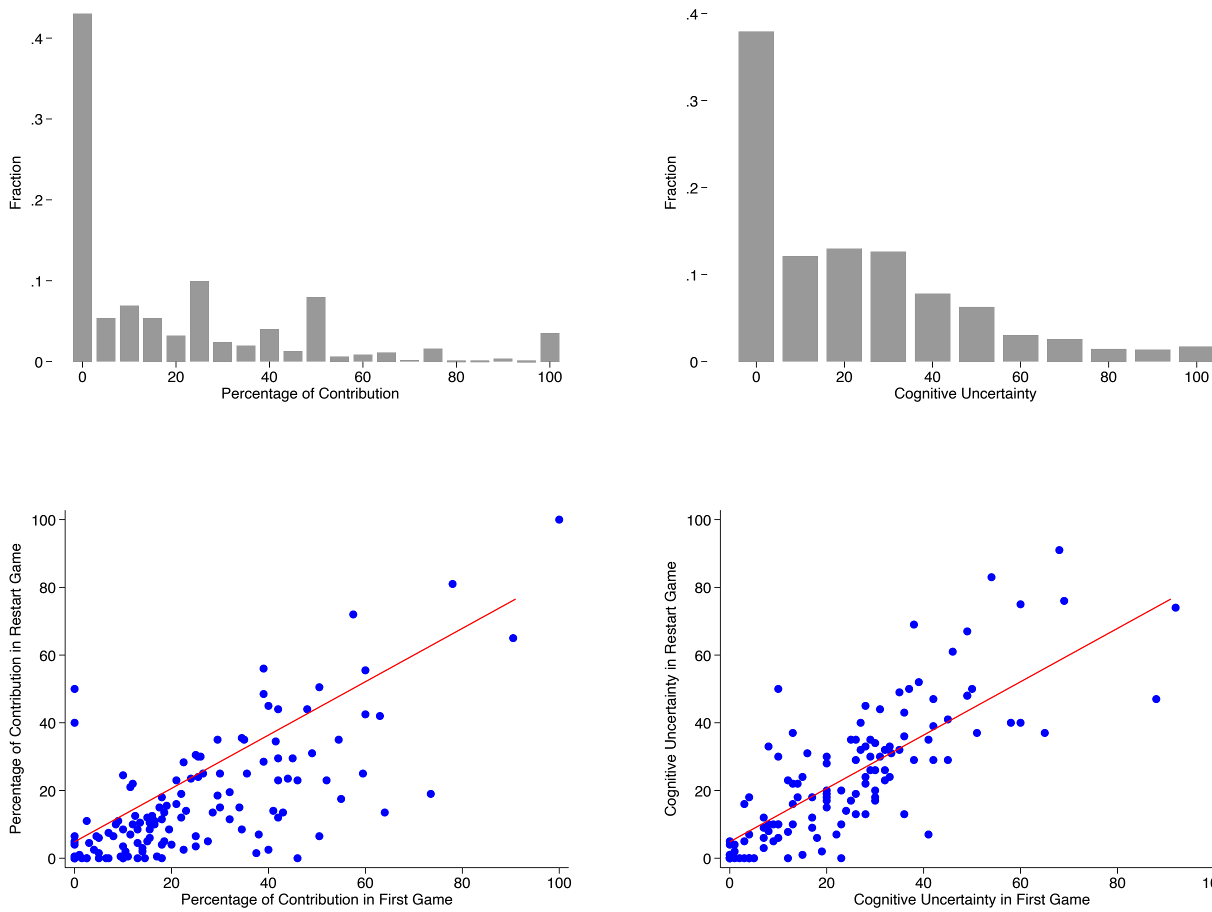



Figure 1. Upper panel: Histogram of percentage of contribution (left panel, $N = 2978$) and cognitive uncertainty (right panel, $N = 2978$). Bottom Panel: Correlation between percentage of contribution in the first game vs. in the restart game (left panel, $N$=140, $\rho$=0.725) and correlation between cognitive uncertainty in the first game vs. in the restart game (right panel, $N$=140, $\rho$=0.831)

The implementation of other groups is as discussed in Treatment. One can find the details in Table B.2, and the screenshot of the instruction in Appendix C. We refer to the first 10 rounds as the "first game" and the last 10 rounds as the "restart". It is important to note that since the treatments of recommendation are imposed only after the end of the first game, all groups in the first game are playing the Treatment Baseline.

At the end of the experiment, the experimenter privately paid the participants using the QR code provided by the participants.

### 2.3 Summary Statistics

The histogram of contributions and cognitive uncertainty can be found in Figure 1, and the detailed summary statistics can be found in Table B.1. On average, subjects contributed 19.43% of their endowment, with 43% of the decisions resulting in zero contribution to the public group account. Additionally, subjects display a moderate level of cognitive certainty in their decisions of contribution compared to in other domain in the literature, with an average cognitive uncertainty of around 20% and 37.9% of decisions associated with zero cognitive uncertainty. Similar to Enke et al. (2023), we observe reasonably high within-domain stability, where subjects exhibit consistency in displaying high or low cognitive uncertainty. Specifically, both contribution decisions ($\rho$=0.7247) and cognitive uncertainty ($\rho$=0.8312) are highly correlated before and after the restart of the game. Consequently, we control for subject fixed effects in all of our analyses. Participants earn an average of S$11.17 for a mean completion time of 1 hour.

# 3 Cognitive Uncertainty and Apparent Social preference

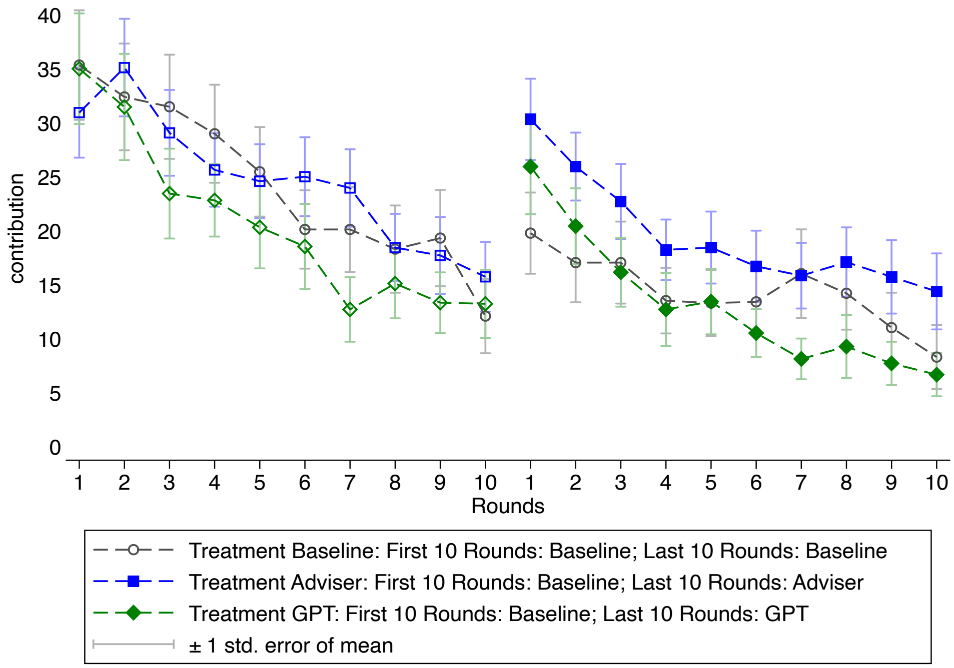
## 3.1 Decreasing Cognitive Uncertainty After Receiving Cooperative Advice

The bottom panel of Figure 2 depicts the evolution of cognitive uncertainty as the game repeats. The main finding is that there is no discernible trend in the evolution of cognitive uncertainty in the Treatment Baseline. In contrast, when subjects are prompted with a recommendation to contribute 65% of their endowment (which is relatively cooperative compared to the average contribution decision of 23% of their endowment in the first game), they become more certain about their contribution decisions as the game repeats. We also plot the evolution of contribution in the upper panel of Figure 2, which shows a similar pattern as the findings in the existing literature (e.g., see Croson, 1996 Figure 1)[6].

Table 1 presents the corresponding aggregate regression estimates that control for subject fixed effects and lagged average contribution from his group members, and separate the subjects by treatment[7]. The results confirm the visual impression from Figure 2. In Treatment Baseline and regardless of it is the first or restart game, we fail to observe a significant trend in the evolution of cognitive uncertainty as the game repeats, as showed in Column (2) and (4)-(6). In contrast, a prompt of cooperative advice increases subjects' cognitive certainty by 11% in total over the 10 rounds of games. The lagged average contribution from group members is not found have significant impact on subjects' cognitive uncertainty. Meanwhile, we conduct Wald Test of the slope of cognitive uncertainty on round, comparing between that in the Restart game against in the First game. There is a significant change in the slope of cognitive uncertainty in Group Adviser ($p$=0.007) and Group GPT ($p$=0.022), but the slope does not show significant change in the Group Baseline (p=0.739). We also fail to find any significant difference in the slope in Treatment GPT

[6]As shown in the bottom panel of Figure 2, participants contribute between 30-40% of their endowment in the first round, but the contribution decays to 10%-20% when the game is repeated. When the game is simply restarted, a restart effect is observed, where there is a slight increment in contribution at the beginning of the restart, followed by a decay in contribution as the game repeats, similar to the pattern observed in the first game.

[7]Table B.3 presents the regression table, similar to Table 1, but with subjects grouped according to their assigned groups instead of the treatment. It displays comparable results to those in Table 1.



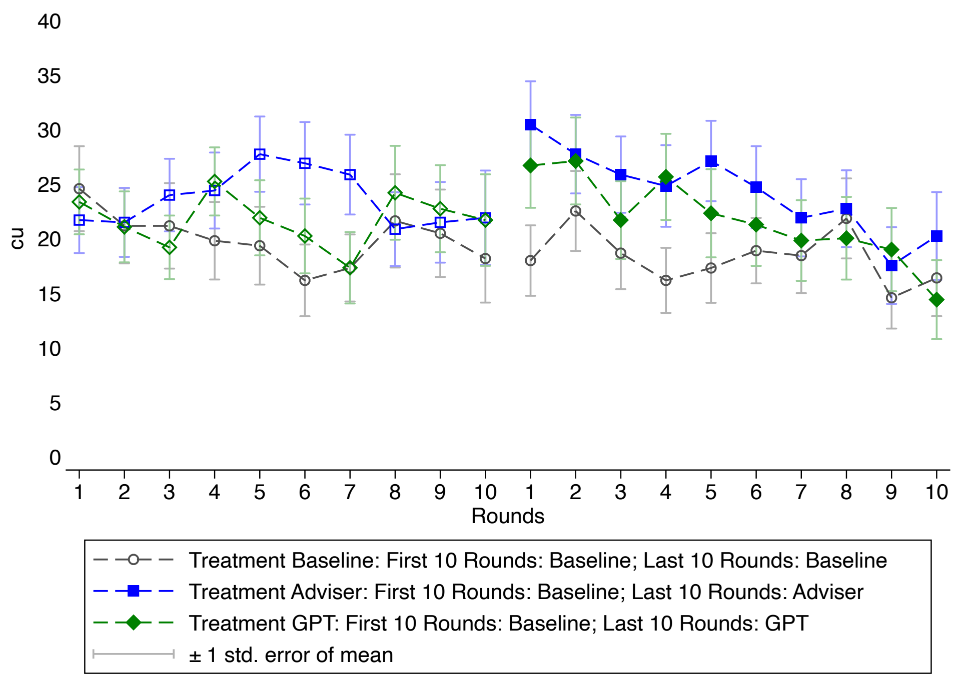




Figure 2. Upper panel: Histogram of percentage of contribution (left panel, $N$ = 2978) and cognitive uncertainty (right panel, $N$ = 2978). Bottom Panel: Correlation between percentage of contribution in the first game vs. in the restart game (left panel, $N$=140, $\rho$=0.725) and correlation between cognitive uncertainty in the first game vs. in the restart game (right panel, $N$=140, $\rho$=0.831)

Table 1. Variation of Cognitive Uncertainty with respect to Rounds

| | Dependent Variable: Cognitive Uncertainty | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Treatment | Pooled | | | Baseline | | | Adviser | GPT |
| Game | All | First | Restart | All | First | Restart | Restart | Restart |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Round | -0.15 | -0.20 | -0.86*** | -0.16 | -0.20 | -0.27 | -1.10*** | -1.21*** |
| | (0.10) | (0.25) | (0.21) | (0.16) | (0.25) | (0.40) | (0.36) | (0.36) |
| Average Lagged group members' % contribution | 0.00 | -0.04 | 0.03 | -0.02 | -0.04 | 0.03 | 0.06 | -0.02 |
| | (0.03) | (0.04) | (0.04) | (0.04) | (0.04) | (0.07) | (0.08) | (0.07) |
| Observations | 2,656 | 1,259 | 1,397 | 1,698 | 1,259 | 439 | 478 | 480 |
| R-squared | 0.60 | 0.63 | 0.67 | 0.59 | 0.63 | 0.59 | 0.66 | 0.73 |

Notes. Subject level fixed effect OLS estimates, with robust standard errors (in parentheses) are clustered at the subject level. Column (1)-(3) pooled all data in all groups and treatment but separate them by the game. Column (4)-(6) include all the data in Treatment Baseline. Specifically, Column (5) include data from all treatment in the First game, while Column (6) only include data from Group Baseline in the Restart Game. Column (7)-(8) restrict attention to the decisions in the Restart Game of Group Adviser and Group GPT, respectively. *** p<0.01, ** p<0.05, * p<0.1.

and Treatment Adviser ($p$=0.922). The results from Wald test are consistent with our previous observation, where a cooperative advice assists individuals to become more cognitively certain about their decisions as the game repeats.

It is important to note that our analysis focuses solely on the evolution of cognitive uncertainty as the game repeats in response to cooperative advice. The analysis on whether the cooperative advice increases or decreases cognitive uncertainty on level will be addressed in Section 3.3.1.

**Result 1:** *Subjects exhibit a decreasing trend in cognitive uncertainty as the game repeats but only after being prompted with cooperative advice from either an anonymous adviser or GPT-3.5.*

The implication our Result 1 is that the dynamics of cognitive uncertainty depend on whether subjects receive cooperative advice before making their decisions. Specifically, in the setting of the classic public goods game and as the game repeats, there is no clear trend in cognitive uncertainty. However, when subjects were provided with cooperative advice before making their decisions, regardless of whether the advice came from an anonymous adviser or GPT, they showed a decreasing cognitive uncertainty in their decisions as the game repeated. Generally speaking, our experimental results demonstrate that a cooperative advice can assist individual in either gaining a better understanding of their true preference or translating their true preferences into contribution actions that maximize their utility as the game repeats. We also find that advice from the GPT reduces cognitive uncertainty for all participants, though the impact of the advice does not seem to depend on whether or not the participants are informed the advice was made by GPT.

### 3.2 Linking Cognitive Uncertainty to Contribution Decision

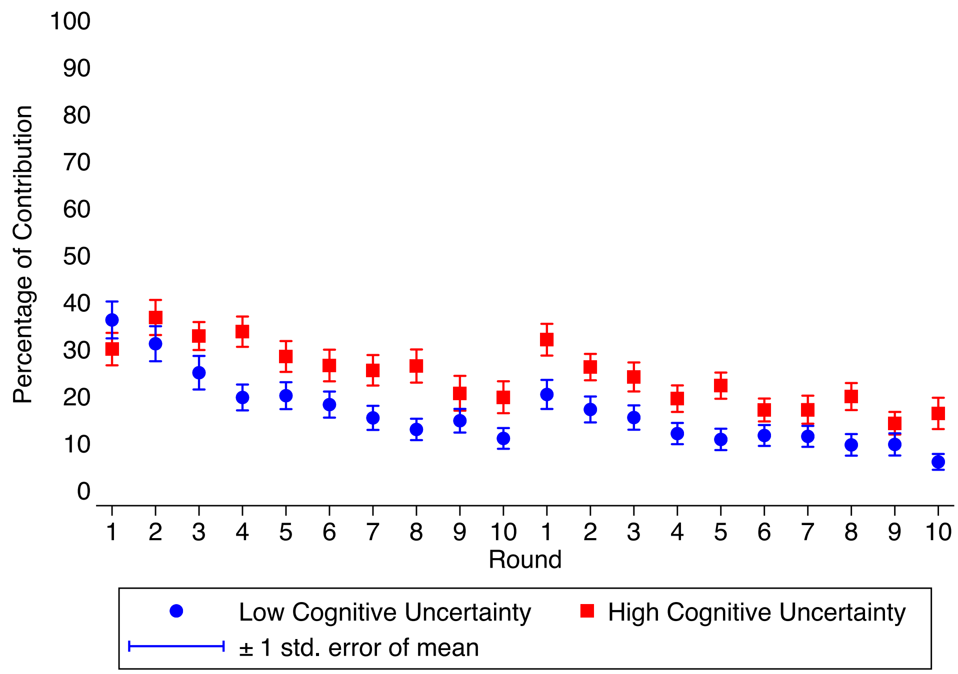



Figure 3. Full Sample. Average contribution percentage to the public goods. Cognitive uncertainty is distinguished by comparing with the average of the cognitive uncertainty pooling all data in the experiment within a given round. High cognitive uncertainty decisions are decisions with cognitive uncertainty that are larger or equal to the average cognitive uncertainty within a given round, while low cognitive uncertainty are those with a cognitive uncertainty that are smaller than the average cognitive uncertainty within a given round.

#### 3.2.1 Cognitively Uncertain Decisions are Closer the Cognitive Default

In this section, we examine how cognitive uncertainty affects contribution decisions. We begin by replicating the graphical analysis conducted in Enke and Graeber (2023). Figure 3 displays the raw data on how contribution decisions evolve across rounds when pooling all the data, distinguishing between decisions associated with above-average cognitive uncertainty (High Cognitive Uncertainty) and below-average cognitive uncertainty (Low Cognitive Uncertainty) within a given round. An analogous figure, where the sample is separated by treatment, can be found in Figure A.1-Figure A.3.

In contrast to the inelastic pattern observed in cognitively uncertain decisions in their study, we do not find any difference in elasticity between decisions with high and low cognitive uncertainty. This visual observation is supported by Table B.4, where we fail to find an opposite and statistically

Table 2. Absolute Deviation from Cognitive Default with regards to Cognitive uncertainty

| | Dependent Variable: Absolute Contribution Deviation from 50% of the Endowment | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Treatment | Pooled | | | Baseline | | | Adviser | GPT |
| Game | All | First | Restart | All | First | Restart | Restart | Restart |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Cognitive Uncertainty | -0.16*** | -0.14*** | -0.18*** | -0.15*** | -0.14*** | -0.20*** | -0.13* | -0.22*** |
| | (0.03) | (0.05) | (0.04) | (0.04) | (0.05) | (0.06) | (0.07) | (0.06) |
| Average Lagged group members' % contribution | -0.20*** | -0.19*** | -0.18*** | -0.20*** | -0.19*** | -0.17*** | -0.19*** | -0.17*** |
| | (0.02) | (0.03) | (0.03) | (0.03) | (0.03) | (0.04) | (0.04) | (0.05) |
| Observations | 2,656 | 1,259 | 1,397 | 1,698 | 1,259 | 439 | 478 | 480 |
| R-squared | 0.12 | 0.10 | 0.12 | 0.10 | 0.10 | 0.14 | 0.09 | 0.16 |
| Number of Subject | 140 | 140 | 140 | 140 | 140 | 44 | 48 | 48 |

Notes. Fixed effects model with cluster-robust standard errors for panels nested within subject level. Column (1)-(3) pooled all data in all groups and treatment but separate them by the game. Column (4)-(6) include all the data in Treatment Baseline. Specifically, Column (5) include data from all treatment in the First game, while Column (6) only include data from Group Baseline in the Restart Game. Column (7)-(8) restrict attention to the decisions in the Restart Game of Group Adviser and Group GPT, respectively. *** p<0.01, ** p<0.05, * p<0.1

significant coefficient between round and the interaction term (between cognitive uncertainty and round).

However, this does not mean that our results do not support their hypothesis suggesting that cognitively uncertain decisions are closer to the heuristic default of 50% contribution. The lack of significant differences in contribution decision patterns can be attributed primarily to the prevalence of the low contribution decisions in our experiment, as indicated by the histogram in Figure 1. This resulted in a scarcity of data points demonstrating a high contribution percentage across various levels of cognitive uncertainty.

As shown in Figure 3 and Figure A.1-A.3, contribution decisions associated with higher cognitive uncertainty tend to be closer to the 50% heuristic compared to those with lower cognitive uncertainty. The results from fixed-effects panel data regression estimates, controlling for payment history in Table 2, confirm the impression conveyed in Figure 3. Specifically, Table 2 demonstrates that cognitively uncertain decisions are significantly closer to the heuristic default of 50% contribution in Column (2). In Table B.5, we replace the dependent variable with the absolute difference from 65% contribution, which corresponds to the recommendation provided to participants in Treatment GPT and Treatment Adviser. Once again, we find that all columns show a negative and statistically significant result, indicating that cognitively uncertain decisions are closer to the cognitive default. We come to our second conclusion.

**Result 2:** Our experimental data support the cognitive uncertainty hypothesis proposed by Enke and Graeber (2023). Using a correlational study, we find that subjects assign less weight to the cognitive default (50% contribution, implying a middle bias) when they are more cognitively certain about their decisions. In turn, the cognitively certain decisions deviate further from the cognitive default and represent more of their true social preference.

### 3.2.2 Cognitive Noise Makes People Contribute More on an Aggregate Level

The cognitive uncertainty model suggests that when people experience cognitive uncertainty, they assign more weight to a default and less weight to their true preference. In this section, we examine how a cognitive noise biases the apparent social preference.

As shown in Figure 2, the overall contribution levels of cognitively uncertain decisions are higher compared to cognitively certain decisions. The results from fixed-effects panel data regression estimates, controlling for lagged average contribution from group members in Table 3, confirm this observation. When decision-makers are completely uncertain about their decisions, they contribute 20% to 30% more than when they are completely cognitively certain. In Appendix E, we conduct a separate causal study. In the treatment Complexity, we increase the complexity by displaying the multiplier in a complicated mathematical equation. We find that both cognitive uncertainty and contribution are higher in the treatment Complexity. In conclusion, both correlational and causal results confirm that at the aggregate level, cognitive uncertainty leads individuals to contribute more to the public account, even though they are less cooperative than their decisions imply[8].

To test how much cognitive uncertainty alone could explain the contribution decision in public goods game, we first tabulate the summary statistics cognitive uncertainty in different level of contribution in Table B.6. There is an increasing trend of cognitive uncertainty when contribution is smaller than 30, but not clear afterwards. We further repeat the analysis in Table 3 Panel A but only include cognitive uncertainty as the sole explanatory variable. The result is showed in Table B.7. When only including cognitive uncertainty into the regression on contribution, the $R^2$ stays at a low level (i.e., between 0.03 and 0.06). Meanwhile, on the aggregate level, a 100% increase in the cognitive uncertainty could only increase contribution by 18% - 25% of the contribution. We also find that when other group members contribute 100% more in the last period on average, it leads subject to contribute at least 18% more in this period.

Furthermore, we conduct a subject-level fixed effect regression analysis of contribution, with additionally control for the confusion (i.e., cognitive uncertainty) and lagged average contributions from other group members, to find out whether the declining contribution by round is due to confusion or frustrated attempts at kindness. The result is showed in Table B.8. The larger

[8]Meanwhile, we conduct Wald Test of the slope of contribution as game repeats, comparing between that in the Restart game against in the First game. There is a significant change in the slope of contribution in Group Adviser (p=0.000) and Group GPT ($p$=0.000), but the slope does not show significant change at 5% in the Group Baseline ($p$=0.083). We also fail to find any significant difference in the slope in Treatment GPT and Treatment Adviser ($p$=0.622). The results from Wald test are consistent with our previous observation on the change of slope in cognitive uncertainty, and further signal that contribution is affected by the cognitive uncertainty.

Table 3. Percentage of Contribution with respect to Cognitive Uncertainty

| | Dependent Variable: Percentage of Contribution | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Treatment | Pooled | | | Baseline | | | Adviser | GPT |
| Game | All | First | Restart | All | First | Restart | Restart | Restart |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| **Panel A: Aggregate Analysis** | | | | | | | | |
| Cognitive Uncertainty | 0.23*** | 0.25*** | 0.21*** | 0.25*** | 0.25*** | 0.24*** | 0.18** | 0.23*** |
| | (0.05) | (0.07) | (0.04) | (0.06) | (0.07) | (0.08) | (0.08) | (0.07) |
| Average Lagged group members' % contribution | 0.30*** | 0.31*** | 0.20*** | 0.31*** | 0.31*** | 0.18*** | 0.21*** | 0.23*** |
| | (0.03) | (0.04) | (0.03) | (0.04) | (0.04) | (0.04) | (0.05) | (0.07) |
| Observations | 2,656 | 1,259 | 1,397 | 1,698 | 1,259 | 439 | 478 | 480 |
| R-squared | 0.12 | 0.12 | 0.09 | 0.12 | 0.12 | 0.10 | 0.10 | 0.08 |
| Number of Subject | 140 | 140 | 140 | 140 | 140 | 44 | 48 | 48 |
| **Panel B: Individual Analysis** | | | | | | | | |
| Cognitive uncertainty makes them contribute more | | | | | | | | |
| No. subjects | 40 | 25 | 27 | 15 | 9 | 9 | 8 | 10 |
| Cognitive uncertainty makes them contribute less | | | | | | | | |
| No. subjects | 5 | 11 | 5 | 2 | 3 | 0 | 2 | 3 |
| Total Subjects | 140 | 140 | 140 | 44 | 44 | 44 | 48 | 48 |

Notes. Fixed effects model with cluster-robust standard errors for panels nested within subject level. Column (1)-(3) pooled all data in all groups and treatment but separate them by the game. Column (4)-(6) include all the data in Treatment Baseline. Specifically, Column (5) include data from all treatment in the First game, while Column (6) only include data from Group Baseline in the Restart Game. Column (7)-(8) restrict attention to the decisions in the Restart Game of Group Adviser and Group GPT, respectively. *** p<0.01, ** p<0.05, * p<0.1. Panel B: OLS estimates on each subject, controlling the lagged average contribution from group members showed in Panel A. Subjects are categorized as cognitive uncertainty make them contribute more (contribute less) when coefficient of cognitive uncertainty on percentage of contribution is greater than (smaller than) zero, and the p-value of the coefficient is less than 0.05.

absolute coefficient on round compared to cognitive uncertainty implies that frustrated attempts at kindness explains more on the declining contribution, compared to confusion, which is consistent with the finding from Andreoni (1995). However, confusion also plays a non-trivial part in it, as they are still statistically significant when control for round.

Overall, our results suggest that cognitive uncertainty cannot explain contribution behavior *alone*. Instead, cognitive uncertainty only complements but not replace the taste-based social preference to explain the contribution decision. Specifically, cognitive noise distorts the contribution decision that primarily reflects subjects' true social preference.

Next, we conduct the individual level analysis to look at if there is any heterogeneity within the subject. i.e., Is there any subject who are in fact more cooperative, while the cognitive uncertainty makes them appear to be less cooperative? The results are presented in Table 3 Panel B. For each subject, we estimate a regression model to analyze how their contribution decisions change with increasing cognitive uncertainty, while controlling for lagged average contribution from group members. Subjects are then categorized as cognitive uncertainty making them contribute more (contribute less) when the coefficient of cognitive uncertainty is greater than (smaller than) zero, and the p-value of the coefficient is less than 0.05.

When pooling all the data from our experiment in Column (1), we find that the majority of subjects (28.6%) significantly contribute more due to cognitive uncertainty, while only a minority of subjects (3.6%) significantly contribute less due to cognitive uncertainty. This pattern remains consistent when separating the data by rounds and treatment. The summary statistics of these two types of subjects are presented in Table B.9. Although the number of subjects where cognitive uncertainty significantly makes them contribute less is small, the effect size is notable. Pooling all the data and among the 5 subjects where cognitive uncertainty significantly makes them contribute less, a 50% increase in cognitive uncertainty would result in a 41.2% reduction in their contribution. The magnitude is even larger than the that of the 40 subjects where cognitive uncertainty significantly makes them contribute more, where a 50% increase in cognitive uncertainty would only result in a 36.3% rise in their contribution[9]. We come to our third conclusion:

**Result 3:** *Cognitive noise complements, rather than replaces, taste-based social preference to explain the contribution decision. We find that cognitive noise distorts the contribution decision that primarily reflects subjects' true social preference. Both correlational and causal data supports the notion that cognitive uncertainty is positively correlated with contribution in the public goods game at the aggregate level, or cognitive uncertainty leads people to behave as if they are more cooperative.*

[9]We also conduct statistical test on the equality of the coefficient of cognitive uncertainty on contribution, where the coefficient is standardized into an absolute value. We find the magnitude of the coefficient of those cognitive uncertainty make them contribute less (from 5 subjects) is larger than that of those cognitive uncertainty make them contribute more (from 40 subjects): Two-sample t-test with unequal variance: $t$=7.6185, $p$=0.001; ranksum: $z$=3.612, Exact $p$=0.0000. Note that the analysis suffers from small sample problem. Meanwhile, the cognitive uncertainty of those whose cognitive uncertainty make them contribute less (from 5 subjects) is statistically significantly larger than those whose cognitive uncertainty make them contribute more (from 40 subjects): Two-sample t-test with unequal variance: $t$=6.4694, $p$=0.000, $N$=899; rank sum: $z$=6.407, Exact $p$=0.0000, $N$=899.

*However, there is heterogeneity, where cognitive noise is negatively correlated with the contribution level some participants at an economically significant extent.*

The implication of Result 3 is that future studies on public goods game should measure cognitive uncertainty and take cognitive uncertainty into consideration. At the aggregate level, cognitive noise leads people to contribute more, whereas in reality, individuals are less cooperative than suggested by their contribution decisions in the public goods game due to cognitive noise. Moreover, our data also reveals heterogeneity in how cognitive uncertainty affects apparent social preference. While the cognitive noise is positively correlated with the contribution level of the majority of the subjects, there is still a minority of subjects whose level of contribution is actually negatively correlated with cognitive noise at an economically significant extent. Therefore, when researchers are interested in determining subjects' true social preference, they should consider their contribution decisions only when the corresponding cognitive uncertainty is below an acceptable threshold.

## 3.3 Impact of Cooperative Advice From Adviser and GPT

In the previous section, we examined how a cooperative advice would impact the change in the slope of cognitive uncertainty with respect to the rounds of the game. We observed that subjects found the contribution decision to be less cognitively uncertain only after being prompted with cooperative advice from either an anonymous adviser or GPT-3.5.

In this section, we investigate how a cooperative advice would affect the level change in participants' cognitive uncertainty and contribution decisions. In other words, we ask whether participants trust an anonymous adviser and GPT-3.5 to a different level?

Since subjects' reactions to the advice given by GPT may be influenced by their prior experience with GPT, we further divide the sample into two groups: those who are inexperienced with GPT (i.e., those who have never heard of GPT, heard of it but never used it, or only used it once prior to the experiment), and those who are experienced with GPT (i.e., those who have used GPT more than once prior to the experiment)[10].

### 3.3.1 Local Effect (RD Analysis)

We begin by focusing on the short-run local effect. Figure 2 presents line graphs depicting the changes in cognitive uncertainty and contribution decisions as the game is repeated. With the exception of cognitive uncertainty in Group Baseline, we observe a jump in both contribution decision and cognitive uncertainty in all other groups at the restart of the game. Figure A.2 and Figure A.3 display regression discontinuity plots at the cutoff point of round = 10.5, separately for

[10] In the Treatment Baseline, there are 20 inexperienced GPT user and 24 experienced GPT user. In Adviser treatment, there are 17 inexperienced GPT user and 31 experienced GPT user. In GPT treatment, there are 20 inexperienced GPT user and 28 experienced GPT user.

Table 4. Covariate-adjusted Sharp RD Estimates using Local Polynomial Regression

| Dependent Variable | Percentage of Contribution | | | | Cognitive Uncertainty | | | |
|---|---|---|---|---|---|---|---|---|
| Group | Pooled (1) | Baseline (2) | Adviser (3) | GPT (4) | Pooled (5) | Baseline (6) | Adviser (7) | GPT (8) |
| **Panel A: Pooled** | | | | | | | | |
| Sharp RD Estimate | 14.49*** | 8.78* | 17.32*** | 14.68** | 5.12 | 0.24 | 10.25* | 4.55 |
| | (3.22) | (5.27) | (4.92) | (5.78) | (3.36) | (5.54) | (5.90) | (5.98) |
| Observations | 2,656 | 789 | 861 | 864 | 2,656 | 789 | 861 | 864 |
| **Panel B: Inexperienced GPT user** | | | | | | | | |
| Sharp RD Estimate | 12.41*** | 6.71 | 17.49*** | 13.96* | 7.36 | 1.75 | 5.72 | 12.99 |
| | (4.01) | (6.54) | (6.12) | (8.10) | (4.83) | (6.68) | (7.83) | (10.18) |
| Observations | 1,083 | 360 | 306 | 360 | 1,083 | 360 | 306 | 360 |
| **Panel C: Experienced GPT user** | | | | | | | | |
| Sharp RD Estimate | 15.97*** | 10.72 | 17.19** | 14.37* | 3.81 | -0.08 | 12.58 | -1.63 |
| | (4.58) | (8.13) | (6.83) | (7.93) | (4.68) | (8.85) | (8.19) | (7.09) |
| Observations | 1,573 | 429 | 555 | 504 | 1,573 | 429 | 555 | 504 |

Notes. Local polynomial sharp regression discontinuity estimates, RD cutoff point at round = 10.5 and additionally control for the lagged average contribution by the group members, with robust standard error clustered (in parentheses) clustered at subject level. *** p<0.01, ** p<0.05, * p<0.1

each group. The graphs capture the local effect of the restart in Group Baseline and the combined effect of the restart and cooperative advice in Group Adviser and Group GPT. The treatment effect appears to be visually larger in Group Adviser and Group GPT compared to Group Baseline in the contribution decision. Meanwhile, the jump is larger in the contribution compared to that in the cognitive uncertainty.

Table 4 presents the corresponding sharp RD estimates. We find a statistically significant local effect of the cooperative advice on contribution, where there are larger magnitudes in Column (3) and (4), compared to Column (2) which only captures only the restart effect. The finding is generally robust to subjects' prior experience with GPT. In contrast, we do not find a statistically significant restart effect or an effect from the cooperative advice on cognitive uncertainty at the 5

### 3.3.2 Global Effect (DID Analysis)

Next, we examine whether there is a sustained effect from cooperative advice on cognitive uncertainty and contribution. We estimate the Average Treatment Effect on the Treated (ATET) after testing if the trend of cognitive uncertainty and contribution is parallel before the restart of the game (where all groups were playing the Treatment Baseline). Additionally, we depict the evolution of contribution as the game is repeated, which is shown in the upper panel of Figure 1. The main takeaway from the figure is that there is no significant difference in the First Game between the three treatments, which is confirmed by the parallel trend tests in Table B.10 - B.13.[11] The estimation results on the treatment effect are presented in these tables, with significant treatment effects of at least a 10% significance level and a parallel trend highlighted with a border.

Overall, we observe a similar pattern to the RD analysis, where there is only a significant and positive effect on contribution from cooperative advice. However, this effect does not appear to persist beyond 4 rounds in the restart game, as shown in Column (1) and (2) of Table B.10 and Table B.11. This positive effect on the contribution decision is sustained and pronounced among those who are experienced with GPT. By contrast, as showed in Table B.12 and Table B.13, the effect on cognitive uncertainty is not sustained and lasts at most for two period after the treatment, which lasts for a much shorter period compared to the effect on contribution. We therefore come to our fourth conclusion:

**Result 4:** *Our experimental results support the hypothesis that subjects' responses may be biased towards GPT's decisions in the absence of supervision during online experiments. However, the impact from GPT is relatively short run and un-sustained for 10 rounds of the game. But in the short run, whether they specifically bias their decisions towards the explanation provided by GPT cannot be simply captured by changes in cognitive uncertainty, at least in the context where GPT advises a cooperative contribution decision. Meanwhile, we did not observe any significant and sustained difference between Group Adviser and Group GPT in both cognitive uncertainty and contribution decision, indicating the absence of a GPT premium.*

Our Result 4 implies that future studies aiming to explore subjects' preferences through online

[11] We run differences-in-differences regression (DID) for panel data to compare if the differences between the two groups change significantly after the respective treatments are imposed. Each unit of observation is a decision made by a participant at a period, with the standard errors clustered at the subject level in the following form: $y=\beta_0+\beta_1\times\text{Round}+\beta_2\times\text{Treat}+\beta_3\times\text{Round}\times\text{Treat}+\epsilon$. Treat = 1 when the round number of treatment group is greater than 10. y represent contribution in Table B.8 and cognitive uncertainty in Table B.9. To look at how long the effect lasts, we keep the sample in the treatment from only Round = 11; Round = 11 and 12; ... Round = 11 to 20 and run the DID regression separately in each table. For each regression, we also conduct the postestimation parallel trend and report the p-value. Specifically, we call a p-value on parallel test that is smaller 0.05 "linear trend not parallel at 95% level", meaning that the two groups are already differ from each other before the treatment is imposed. We also report the coefficient and robust standard error of ATET. As a result, we call a valid treatment effect when p-value of ATET is smaller than 0.1, with the p-value of parallel trend greater than 0.05, which are labeled with a border. A sustained effect would therefore be concluded when the abovementioned effect (represented by a bordered cell) is sustained when we include more periods after the treatment into the sample. By contrast, a short run effect would be concluded when the abovementioned effect (represented by a bordered cell) lasts only several periods.

experiments will need to conduct additional lab experiments for robustness. Further, the lack of GPT premium provides a useful perspective on how AI-human interaction influences human decision on the social context.

## 4 Conclusion

This paper examines the link between cognitive noise and contribution decisions in the public goods game.

Our results demonstrate that a cooperative advice can assist individual in either gaining a better understanding of their true preference or translating their true preferences into contribution actions that maximize their utility as the game repeats. We also find that GPT advice reduces cognitive uncertainty for all participants, though the impact of the advice does not seem to depend on whether or not the participants are informed the advice was made by GPT.

Further, we argue that cognitive noise complements, rather than replaces, taste-based social preference to explain the contribution decision. Specifically, cognitive noise distorts the contribution decision that primarily reflects subjects' true social preference. Across all treatments, we observe that subjects assign more weight to the cognitive default (50% contribution of their endowment, implying a middle bias) when they are less cognitively certain about their decisions. Therefore, both correlational and causal data from our experiment support the cognitive uncertainty hypothesis proposed by Enke and Graeber (2023), and the result is robust when removing strategic uncertainty.

At the aggregate level, cognitive uncertainty leads individuals to contribute more to the public account, which means that they may be less cooperative than what their decisions imply. We also find heterogeneity among the participants in terms of the direction in which cognitive noise biases their apparent social preference. While cognitive noise is positively correlated with the contribution level of the majority of subjects, there is still a minority of subjects whose level of contribution is negatively correlated with cognitive noise to the extent that is economically significant. In other words, there are a minority of subjects in our experiment are in fact more cooperative, while cognitive noise makes them appear to be less cooperative. These results suggest that when researchers are interested in determining subjects' true social preference, they should consider their contribution decisions only when the corresponding cognitive uncertainty is below an acceptable threshold.

Meanwhile, our experimental results support the hypothesis that subjects' responses may be biased towards GPT's decisions in the absence of supervision during online experiments. Furthermore, whether they specifically bias their decisions towards the explanation provided by GPT cannot be simply captured by changes in cognitive uncertainty, at least in the context where GPT advises a cooperative contribution decision. These findings suggest that future studies aiming to explore subjects' preferences through online experiments will need to conduct additional supervised laboratory experiments for robustness. Meanwhile, we did not observe any significant and sustained difference between Group Adviser and Group GPT in both cognitive uncertainty and contribution decision, indicating the absence of a GPT premium. Our result provides a useful perspective on how AI-human interaction influences human decision on the social context.

## References


Amelio, A. (2022). Cognitive uncertainty and overconfidence. Technical report, ECONtribute Discussion Paper.

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

Au, W. (2004). Criticality and environmental uncertainty in step-level public goods dilemmas. *Group Dynamics: Theory, Research, and Practice*, 8(1):40.

Bai, L., Wei, L., and Xue, L. (2023). Endogenous ai-tocracy. *Available at SSRN 4636843*.

Bao, T., Liang, B., and Pei, J. (2022). Does ethnic diversity always undermine pro-social behavior? evidence from a laboratory experiment. *European Journal of Political Economy*, 72:102119.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901.

Burton-Chellew, M. N. (2022). The restart effect in social dilemmas shows humans are self-interested not altruistic. *Proceedings of the National Academy of Sciences*, 119(49):e2210082119.

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

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

Chen, D. L., Schonger, M., and Wickens, C. (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Chen, M., Tworek, J., Jun, H., Yuan, Q., Pinto, H. P. d. O., Kaplan, J., Edwards, H., Burda, Y., Joseph, N., Brockman, G., et al. (2021). Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374*.

Chen, Y., Liu, T. X., Shan, Y., and Zhong, S. (2023). The emergence of economic rationality of gpt. *Proceedings of the National Academy of Sciences*, 120(51):e2316205120.

Cone, J. and Rand, D. G. (2014). Time pressure increases cooperation in competitively framed social dilemmas. *PLoS one*, 9(12):e115756.

Cornelissen, G., Dewitte, S., and Warlop, L. (2011). Are social value orientations expressed automatically? decision making in the dictator game. *Personality and Social Psychology Bulletin*, 37(8):1080–1090.

Croson, R. T. (1996). Partners and strangers revisited. *Economics Letters*, 53(1):25–32.

Drori, I., Zhang, S., Shuttleworth, R., Tang, L., Lu, A., Ke, E., Liu, K., Chen, L., Tran, S., Cheng, N., et al. (2022). A neural network solves, explains, and generates university math problems by program synthesis and few-shot learning at human level. *Proceedings of the National Academy of Sciences*, 119(32):e2123433119.

Enke, B. and Graeber, T. (2023). Cognitive uncertainty. *The Quarterly Journal of Economics*, 138(4):2021–2067.

Enke, B., Graeber, T., and Oprea, R. (2023). Complexity and hyperbolic discounting.

Fehr, E. and Fischbacher, U. (2002). Why social preferences matter–the impact of non-selfish motives on competition, cooperation and incentives. *The economic journal*, 112(478):C1–C33.

Fehr, E. and Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425(6960):785–791.

Fehr, E. and Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4):980–994.

Fischbacher, U. and Gächter, S. (2006). Heterogeneous social preferences and the dynamics of free riding in public goods.

Gangadharan, L. and Nemes, V. (2009). Experimental analysis of risk and uncertainty in provisioning private and public goods. *Economic Inquiry*, 47(1):146–164.

Guo, F. (2023). Gpt in game theory experiments. *arXiv preprint arXiv:2305.05516*.

Isaac, R. M. and Walker, J. M. (1998). Nash as an organizing principle in the voluntary provision of public goods: Experimental evidence. *Experimental economics*, 1:191–206.

Isaac, R. M., Walker, J. M., and Thomas, S. H. (1984). Divergent evidence on free riding: An experimental examination of possible explanations. *Public choice*, 43:113–149.

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

Kappes, A., Nussberger, A.-M., Faber, N. S., Kahane, G., Savulescu, J., and Crockett, M. J. (2018). Uncertainty about the impact of social decisions increases prosocial behaviour. *Nature human behaviour*, 2(8):573–580.

Lin, Z., Xu, P., Winata, G. I., Siddique, F. B., Liu, Z., Shin, J., and Fung, P. (2020). Caire: An end-to-end empathetic chatbot. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 13622–13623.

Messick, D. M., Allison, S. T., and Samuelson, C. D. (1988). Framing and communication effects on group members' responses to environmental and social uncertainty. *Applied behavioral economics*, 2:677–700.

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

Rand, D. G. and Kraft-Todd, G. T. (2014). Reflection does not undermine self-interested prosociality. *Frontiers in behavioral neuroscience*, 8:300.

Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., and Greene, J. D. (2014). Social heuristics shape intuitive cooperation. *Nature communications*, 5(1):3677.

Rapoport, A. and Eshed-Levy, D. (1989). Provision of step-level public goods: Effects of greed and fear of being gypped. *Organizational Behavior and Human Decision Processes*, 44(3):325–344.

Roch, S. G., Lane, J. A., Samuelson, C. D., Allison, S. T., and Dent, J. L. (2000). 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.

Schulz, J. F., Fischbacher, U., Thöni, C., and Utikal, V. (2014). Affect and fairness: Dictator games under cognitive load. *Journal of Economic Psychology*, 41:77–87.

Wang, G., Li, J., Wang, W., Niu, X., and Wang, Y. (2024). Confusion cannot explain cooperative behavior in public goods games. *Proceedings of the National Academy of Sciences*, 121(10):e2310109121.

Wit, A. and Wilke, H. (1998). Public good provision under environmental and social uncertainty. *European journal of social psychology*, 28(2):249–256.

# Appendices

## Appendix A  Additional Figures

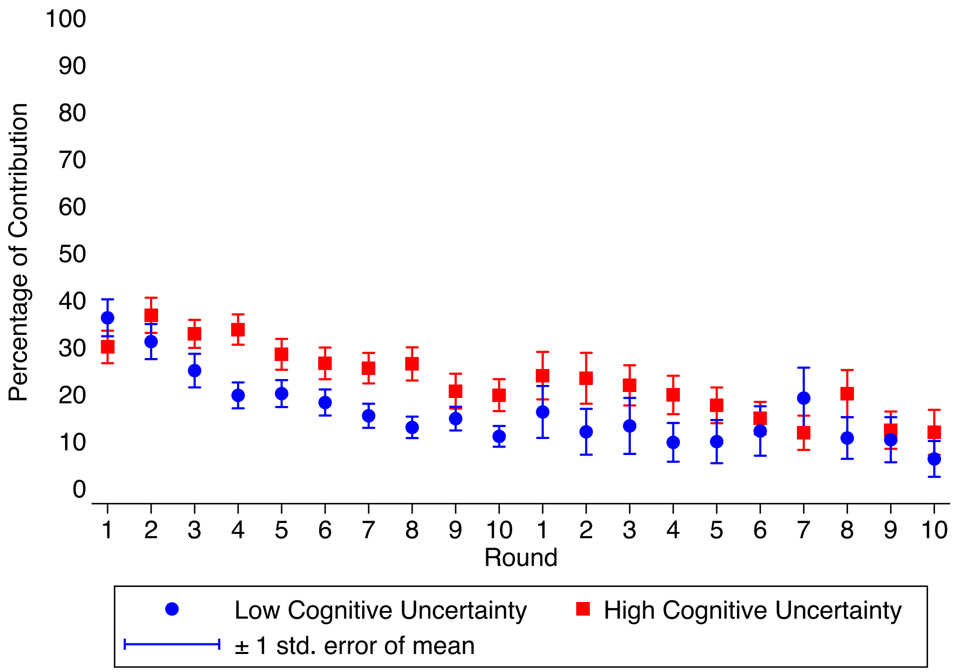



Figure A.1. Treatment Baseline. Average contribution percentage to the public good separating the treatment. Cognitive uncertainty is distinguished by comparing with the average of the cognitive uncertainty in the treatment of the experiment within a given round.

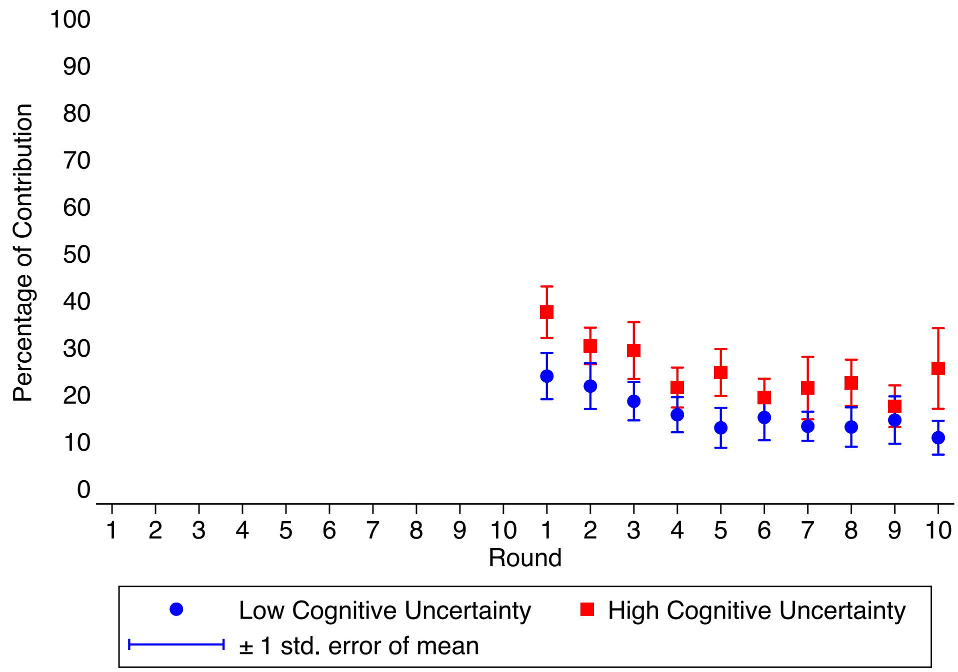



Figure A.2. Treatment Adviser. Average contribution percentage to the public good separating the treatment. Cognitive uncertainty is distinguished by comparing with the average of the cognitive uncertainty in the treatment of the experiment within a given round.

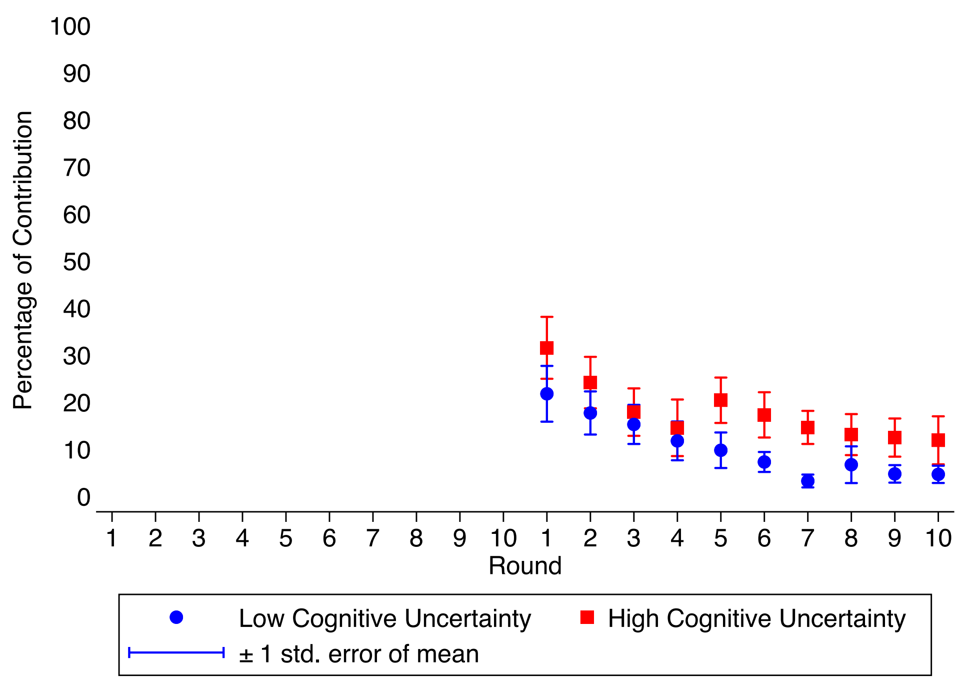



Figure A.3. Treatment GPT. Average contribution percentage to the public good separating the treatment. Cognitive uncertainty is distinguished by comparing with the average of the cognitive uncertainty in the treatment of the experiment within a given round.

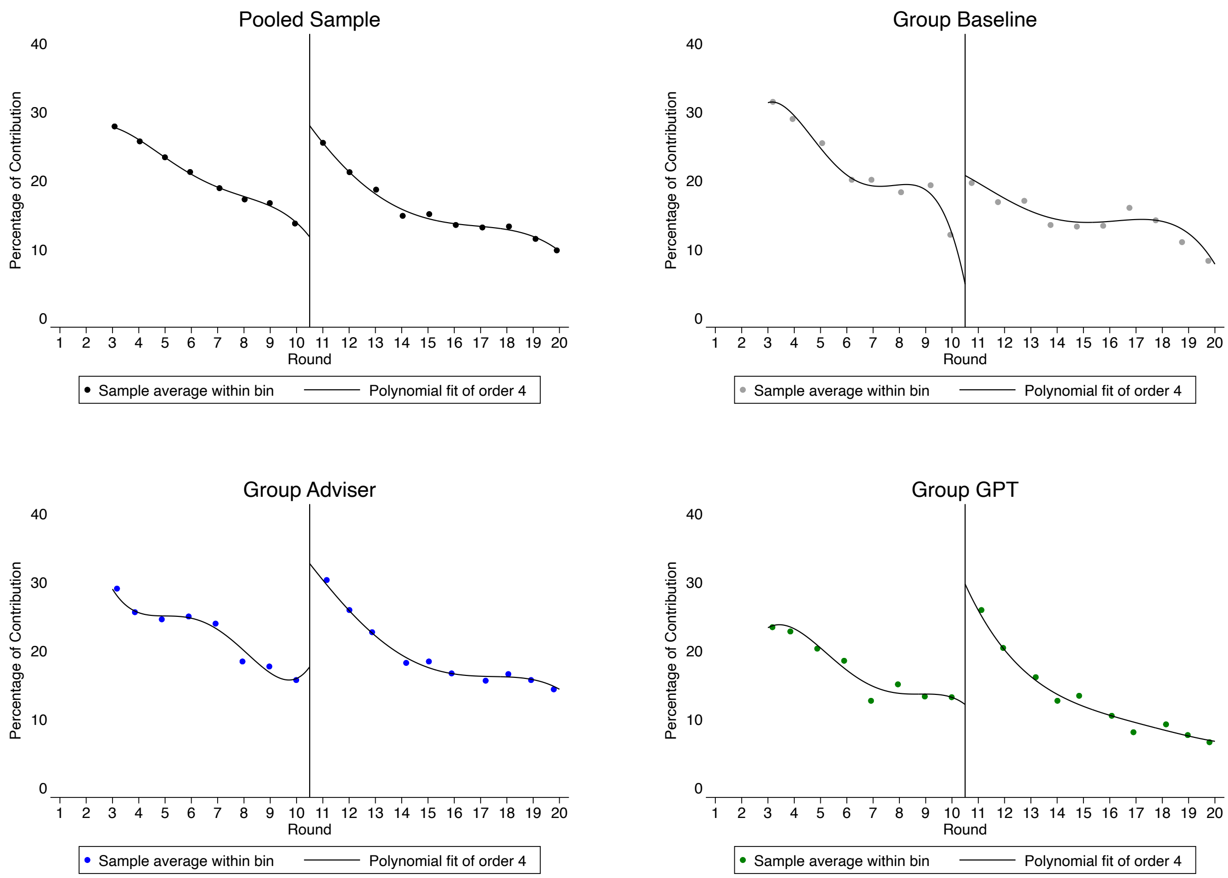



Figure A.4. Regression discontinuity plot on percentage of contribution with evenly spaced mimicking variance number of bins using polynomial regression. Cut-off point at round = 10.5 and additionally control for lagged average contribution from group members.

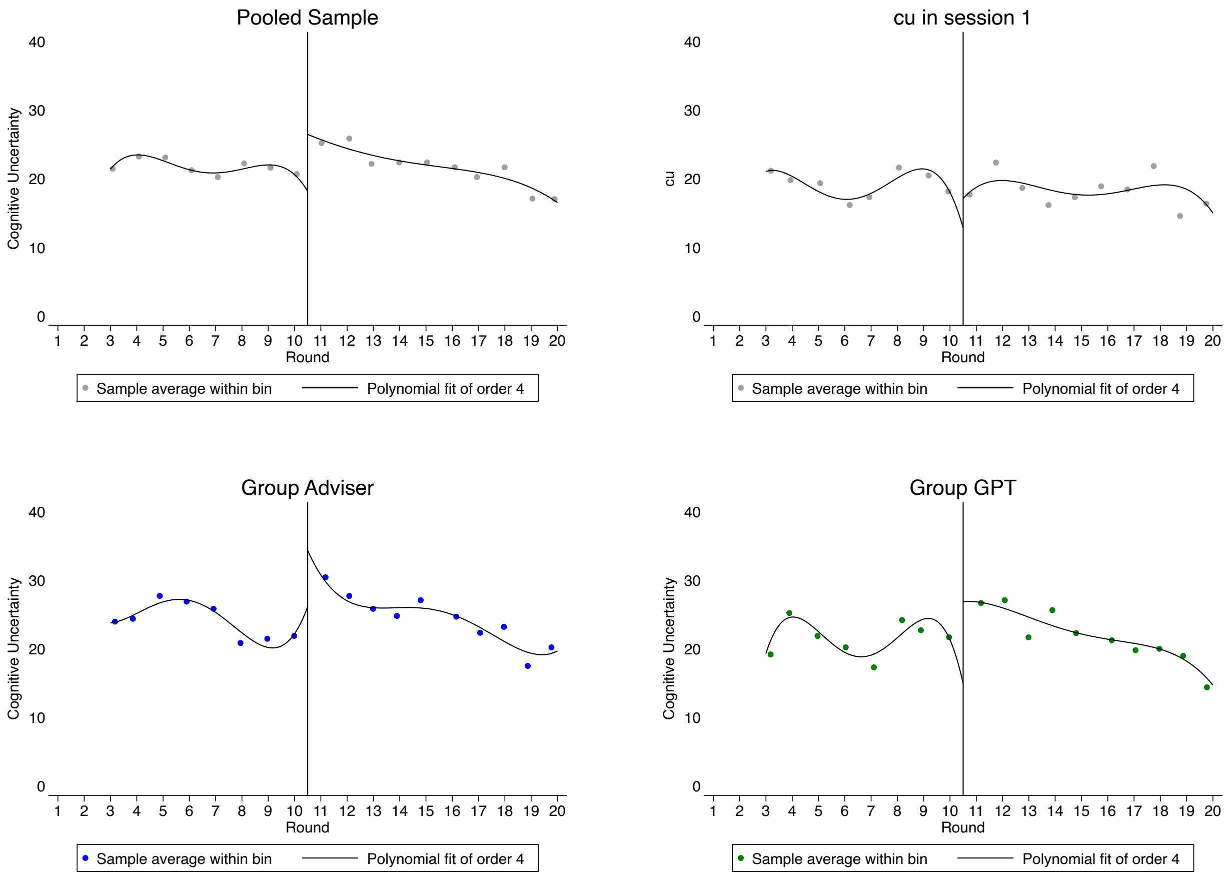



Figure A.5. Regression discontinuity plot on cognitive uncertainty with evenly spaced mimicking variance number of bins using polynomial regression. Cut-off point at round = 10.5 and additionally control for last period payoff.

## Appendix B Additional Tables

Table B.1. Summary Statistics

| | Observation | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|---|
| **Panel A: Full Sample** | | | | | |
| Subject ID | 2,800 | 70.50 | 40.42 | 1 | 140 |
| % Contribution | 2,798 | 19.43 | 25.43 | 0 | 100 |
| Cognitive Uncertainty | 2,798 | 21.66 | 24.64 | 0 | 100 |
| Completion Time (in Min) | 2,798 | 58.66 | 17.52 | 41.85 | 84.15 |
| Payoff Per Round (in Points) | 2,798 | 22.34 | 4.562 | 8 | 40 |
| **Panel B: Group Baseline** | | | | | |
| Subject ID | 880 | 22.50 | 12.71 | 1 | 44 |
| % Contribution | 879 | 19.37 | 26.71 | 0 | 100 |
| Cognitive Uncertainty | 879 | 19.10 | 23.17 | 0 | 100 |
| Completion Time (in Min) | 879 | 47.25 | 0.669 | 45.22 | 48.33 |
| Payoff Per Round (in Points) | 879 | 22.31 | 4.751 | 8 | 39 |
| **Panel C: Group Adviser** | | | | | |
| Subject ID | 960 | 68.50 | 13.86 | 45 | 92 |
| % Contribution | 959 | 22.08 | 24.75 | 0 | 100 |
| Cognitive Uncertainty | 959 | 23.94 | 25.00 | 0 | 100 |
| Completion Time (in Min) | 959 | 82.85 | 1.008 | 79.08 | 84.15 |
| Payoff Per Round (in Points) | 959 | 22.69 | 4.504 | 10 | 40 |
| **Panel D: Group GPT** | | | | | |
| Subject ID | 960 | 116.5 | 13.86 | 93 | 140 |
| % Contribution | 960 | 16.84 | 24.66 | 0 | 100 |
| Cognitive Uncertainty | 960 | 21.72 | 25.37 | 0 | 100 |
| Completion Time (in Min) | 960 | 44.93 | 0.843 | 41.85 | 46.18 |
| Payoff Per Round (in Points) | 960 | 22.02 | 4.420 | 8 | 38 |

Note: The two missing values in contribution and cognitive uncertainty decisions are due to participants being temporarily absent during the experiment. The experimenter had to proceed with the experiment as those participants who did not submit their decisions were unable to be contacted for a long time. Additionally, it is noticeable that the Group Adviser took much longer time than the other two treatments. This is because we conducted the Group Adviser as the first session. Since it was a group experiment, subjects could only proceed once all participants had submitted their decisions. However, during the experiment, many subjects did not realize that they need to refresh the page, resulting in prolonged waiting times for each period that were not caused by other participants' delays. In response, the experimenter reminded the subjects to refresh the page when the webpage displaying the message "waiting for other participants" for too long.

Table B.2. List of Treatment Condition

| **Group, abbrev.** | Description | **Treatment** All Rounds | First Game Round 1-10 | Restart Round 11-20 | Sample |
|---|---|---|---|---|---|
| **Baseline** (SPCU_B) | Subjects are placed into groups of four and play for 20 rounds, with a restart prior to Rounds 11. | Baseline | Baseline | Baseline | 44 |
| **Complexity Number,** (SPCU_TB) | There are only 10 rounds of the game, and there is a changing multiplier factor in each round. | Complexity Number | Complexity Number | NA | 48 |
| **Complexity Equation,** (SPCU_T) | Same as Complexity Baseline, except the multiplier factors are displayed in a mathematical function, and subjects only have 25 seconds to make the decision. | Complexity Equation | Complexity Equation | NA | 48 |
| **Complexity Equation Soft Timeout,** (SPCU_ST) | Same as Complexity Equation, except that subjects were given a non-binding time limit of 25 seconds to submit their decisions. | Complexity Equation Soft Timeout | Complexity Equation Soft Timeout | NA | 44 |
| **Robustness** (SPCU_R) | Same as Baseline, except that ¼ of the participants are randomly assign a subset of subjects into the "Full Information" Treatment, so that they will always know the contribution of all other group members before making their own decision. Meanwhile, they will always be paired with three other subjects who do are not in the "Full Information" Treatment, and the other subjects do not know that they are playing with a subject who can access to their contribution decision. | Robustness | Robustness | NA | 84 |
| **GPT** (SPCU_G) | Same as Baseline, except that participants are prompted with a recommendation from GPT-3.5 of contributing 65% of the endowment after the end of the restart game. | Mixed | Baseline | GPT | 48 |
| **Adviser** (SPCU_A) | Same as GPT, except that participants are not told the recommendation is made by GPT-3.5, and only know that the recommendation is from an anonymous adviser. | Mixed | Baseline | Adviser | 48 |

Note: In situations where subjects exceed their decision time in complexity equation treatment or have not made a decision for an extended period in all other treatments, prompting the experimenter to proceed, their decision will be labelled as 'Timeout = Yes. In these cases, they will not earn any money for this period but will instead make a random contribution between 0 and 20 (i.e., their endowment) for the purpose of calculating the group project's return. They will not know that they made a random contribution but will only see their earning is 0 in the payoff page. Other subjects will not be aware that one group member's contribution is generated in this manner, as all participants within the entire session will be unable to view their results until all subjects have finalized their contribution decisions. The contribution decisions and 0 payoff when time out happens will be excluded from our analysis, Meanwhile, they would also only have 10 seconds to view the result, preventing them to deduct what has happened.

Table B.3. Variation of Cognitive Uncertainty with respect to Rounds

| | Dependent Variable: Cognitive Uncertainty | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Rounds: | All Rounds (i.e., Round 1-20) | | | | First Game (Round 1- 10) | | | | Restart (Round 11-20) | | | |
| Group: | Pooled | Baseline | Adviser | GPT | Pooled | Baseline | Adviser | GPT | Pooled | Baseline | Adviser | GPT |
| Treatment | Mixed | Baseline | Mixed | Mixed | Baseline | Baseline | Baseline | Baseline | Mixed | Baseline | Adviser | GPT |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| **Panel A: Aggregate Analysis** | | | | | | | | | | | | |
| Round | -0.15 | -0.13 | -0.15 | -0.16 | -0.20 | -0.20 | -0.35 | -0.06 | -0.86*** | -0.27 | -1.10*** | -1.21*** |
| | (0.10) | (0.19) | (0.17) | (0.17) | (0.25) | (0.42) | (0.47) | (0.44) | (0.21) | (0.40) | (0.36) | (0.36) |
| Average Lagged group members' % contribution | 0.00 | 0.02 | 0.00 | -0.02 | -0.04 | -0.01 | -0.06 | -0.06 | 0.03 | 0.03 | 0.06 | -0.02 |
| | (0.03) | (0.06) | (0.06) | (0.04) | (0.04) | (0.07) | (0.07) | (0.05) | (0.04) | (0.07) | (0.08) | (0.07) |
| Observations | 2,656 | 834 | 910 | 912 | 1,259 | 395 | 432 | 432 | 1,397 | 439 | 478 | 480 |
| R-squared | 0.60 | 0.54 | 0.55 | 0.67 | 0.63 | 0.63 | 0.55 | 0.71 | 0.67 | 0.59 | 0.66 | 0.73 |
| **Panel B: Individual Analysis** | | | | | | | | | | | | |
| Decreasing trend of cognitive uncertainty as game repeats | | | | | | | | | | | | |
| No. subjects | 19 | 5 | 7 | 7 | 16 | 4 | 6 | 6 | 16 | 1 | 8 | 7 |
| Increasing trend of cognitive uncertainty as game repeats | | | | | | | | | | | | |
| No. subjects | 16 | 7 | 5 | 4 | 9 | 2 | 3 | 4 | 8 | 4 | 2 | 2 |
| Total Subjects | 140 | 44 | 48 | 48 | 140 | 44 | 48 | 48 | 140 | 44 | 48 | 48 |

Notes. Panel A: Subject level fixed effect OLS estimates, with robust standard errors (in parentheses) are clustered at the subject level, controlling for lagged average contribution from group members. Column (1)-(4) include data from all rounds of the games. Column (5)-(8) restrict attention to decisions in the first 10 rounds of the game, where no treatment has been imposed in any of the group. By contrast, column (9)-(12) look at the decisions after the restart of the game, where treatment was imposed at the beginning of the new 10 rounds. *** $p<0.01$, ** $p<0.05$, * $p<0.1$; Panel B: OLS estimates on each subject, controlling for the lagged average group members contribution showed in Panel A. Subjects are categorized as showing decreasing (increasing) trend of cognitive uncertainty as game repeats when the coefficient of round is smaller than (greater than) zero, and the p-value of the coefficient is less than 0.05.

Table B.4. Elasticity to Round and Cognitive Uncertainty

| | Dependent Variable: Percentage of Contribution | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Treatment | | Pooled | | | Baseline | | Adviser | GPT |
| Game | All | First | Restart | All | First | Restart | Restart | Restart |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Round | -0.57*** | -1.59*** | -0.95*** | -0.84*** | -1.59*** | -0.29 | -1.06** | -1.56*** |
| | (0.14) | (0.38) | (0.23) | (0.25) | (0.38) | (0.32) | (0.41) | (0.44) |
| Round × Cognitive uncertainty | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.03* | -0.01 | 0.00 |
| | (0.00) | (0.01) | (0.01) | (0.01) | (0.01) | (0.02) | (0.01) | (0.01) |
| Cognitive uncertainty | 0.29*** | 0.27** | 0.33** | 0.28*** | 0.27** | 0.66** | 0.24 | 0.16 |
| | (0.07) | (0.11) | (0.14) | (0.10) | (0.11) | (0.28) | (0.24) | (0.21) |
| Average Lagged group members' % contribution | 0.21*** | 0.20*** | 0.15*** | 0.21*** | 0.20*** | 0.17*** | 0.16*** | 0.11 |
| | (0.03) | (0.04) | (0.04) | (0.04) | (0.04) | (0.05) | (0.06) | (0.07) |
| Observations | 2,656 | 1,259 | 1,397 | 1,698 | 1,259 | 439 | 478 | 480 |
| R-squared | 0.59 | 0.64 | 0.68 | 0.60 | 0.64 | 0.70 | 0.76 | 0.56 |

Notes. Subject level fixed effect OLS estimates, with robust standard errors (in parentheses) are clustered at the subject level, controlling for lagged average contribution from group members . Column (1)-(3) pooled all data in all groups and treatment but separate them by the game. Column (4)-(6) include all the data in Treatment Baseline. Specifically, Column (5) include data from all treatment in the First game, while Column (6) only include data from Group Baseline in the Restart Game. Column (7)-(8) restrict attention to the decisions in the Restart Game of Group Adviser and Group GPT, respectively. *** p<0.01, ** p<0.05, * p<0.1

Table B.5. Absolute Deviation from Cognitive Default with regards to Cognitive Uncertainty

| | Dependent Variable: Absolute Contribution Deviation from 65% of the Endowment | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Treatment | Pooled | | | Baseline | | | Adviser | GPT |
| Game | All | First | Restart | All | First | Restart | Restart | Restart |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Cognitive Uncertainty | -0.20*** | -0.18*** | -0.21*** | -0.18*** | -0.18*** | -0.22*** | -0.17** | -0.24*** |
| | (0.04) | (0.04) | (0.04) | (0.04) | (0.04) | (0.07) | (0.08) | (0.06) |
| Average Lagged group members' % contribution | -0.24*** | -0.23*** | -0.19*** | -0.24*** | -0.23*** | -0.18*** | -0.20*** | -0.19*** |
| | (0.03) | (0.03) | (0.03) | (0.03) | (0.03) | (0.04) | (0.05) | (0.06) |
| Observations | 2,656 | 1,259 | 1,397 | 1,698 | 1,259 | 439 | 478 | 480 |
| R-squared | 0.14 | 0.14 | 0.13 | 0.14 | 0.14 | 0.14 | 0.10 | 0.15 |
| Number of Subject | 140 | 140 | 140 | 140 | 140 | 44 | 48 | 48 |

Notes. Fixed effects model with cluster-robust standard errors for panels nested within subject level, controlling for lagged average contribution from group members. Column (1)-(3) pooled all data in all groups and treatment but separate them by the game. Column (4)-(6) include all the data in Treatment Baseline. Specifically, Column (5) include data from all treatment in the First game, while Column (6) only include data from Group Baseline in the Restart Game. Column (7)-(8) restrict attention to the decisions in the Restart Game of Group Adviser and Group GPT, respectively. *** p<0.01, ** p<0.05, * p<0.1

Table B.6. Summary Statistics of Cognitive Uncertainty by Contribution Level

| Percentage of Contribution | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| 0 | 1204 | 15.174 | 23.056 | 0 | 100 |
| 5 | 150 | 20.267 | 26.520 | 0 | 100 |
| 10 | 193 | 23.938 | 26.907 | 0 | 100 |
| 15 | 151 | 26.887 | 22.066 | 0 | 100 |
| 20 | 90 | 27.222 | 21.931 | 0 | 90 |
| 25 | 278 | 26.691 | 22.219 | 0 | 100 |
| 30 | 68 | 35.294 | 27.179 | 0 | 100 |
| 35 | 55 | 32.545 | 20.836 | 0 | 80 |
| 40 | 112 | 33.125 | 20.622 | 0 | 100 |
| 45 | 36 | 25.278 | 23.602 | 0 | 90 |
| 50 | 223 | 27.04 | 25.009 | 0 | 100 |
| 55 | 18 | 23.333 | 14.142 | 0 | 50 |
| 60 | 24 | 25.833 | 15.857 | 0 | 70 |
| 65 | 32 | 39.062 | 36.576 | 0 | 100 |
| 70 | 5 | 50 | 23.452 | 30 | 90 |
| 75 | 45 | 30.222 | 22.104 | 0 | 80 |
| 80 | 2 | 0 | 0.000 | 0 | 0 |
| 85 | 2 | 25 | 21.213 | 10 | 40 |
| 90 | 10 | 26 | 22.706 | 0 | 70 |
| 95 | 1 | 20 | 0 | 20 | 20 |
| 100 | 99 | 16.869 | 27.945 | 0 | 100 |

Table B.7. Percentage of Contribution with respect to Cognitive Uncertainty

| | Dependent Variable: Percentage of Contribution | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Treatment | Pooled | | | Baseline | | | Adviser | GPT |
| Game | All | First | Restart | All | First | Restart | Restart | Restart |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Cognitive Uncertainty | 0.23*** | 0.21** | 0.22*** | 0.22*** | 0.21** | 0.24*** | 0.20** | 0.24*** |
| | (0.05) | (0.08) | (0.05) | (0.07) | (0.08) | (0.08) | (0.09) | (0.07) |
| Observations | 2,798 | 1,399 | 1,399 | 1,839 | 1,399 | 440 | 479 | 480 |
| R-squared | 0.04 | 0.03 | 0.06 | 0.03 | 0.03 | 0.06 | 0.06 | 0.05 |
| Number of Subject | 140 | 140 | 140 | 140 | 140 | 44 | 48 | 48 |

Notes. Fixed effects model with cluster-robust standard errors for panels nested within subject level. Column (1)-(3) pooled all data in all groups and treatment but separate them by the game. Column (4)-(6) include all the data in Treatment Baseline. Specifically, Column (5) include data from all treatment in the First game, while Column (6) only include data from Group Baseline in the Restart Game. Column (7)-(8) restrict attention to the decisions in the Restart Game of Group Adviser and Group GPT, respectively. *** p<0.01, ** p<0.05, * p<0.1.

Table B.8. Is Declining Contribution due to Confusion or Frustrated Attempts at Kindness?

| | Dependent Variable: Percentage of Contribution | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Treatment | Pooled | | | Baseline | | | Adviser | GPT |
| Game | All | First | Restart | All | First | Restart | Restart | Restart |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Round | -0.69*** | -1.70*** | -1.16*** | -0.94*** | -1.70*** | -0.79** | -1.22*** | -1.55*** |
| | (0.11) | (0.29) | (0.18) | (0.22) | (0.29) | (0.30) | (0.30) | (0.34) |
| Cognitive Uncertainty | 0.22*** | 0.24*** | 0.18*** | 0.24*** | 0.24*** | 0.23*** | 0.13* | 0.16** |
| | (0.04) | (0.07) | (0.04) | (0.06) | (0.07) | (0.07) | (0.08) | (0.06) |
| Average Lagged group members' % contribution | 0.21*** | 0.20*** | 0.15*** | 0.22*** | 0.20*** | 0.17*** | 0.16*** | 0.11 |
| | (0.03) | (0.04) | (0.04) | (0.04) | (0.04) | (0.05) | (0.06) | (0.07) |
| Observations | 2,656 | 1,259 | 1,397 | 1,698 | 1,259 | 439 | 478 | 480 |
| R-squared | 0.59 | 0.64 | 0.68 | 0.60 | 0.64 | 0.70 | 0.76 | 0.56 |

Notes. Fixed effects model with cluster-robust standard errors for panels nested within subject level. Column (1)-(3) pooled all data in all groups and treatment but separate them by the game. Column (4)-(6) include all the data in Treatment Baseline. Specifically, Column (5) include data from all treatment in the First game, while Column (6) only include data from Group Baseline in the Restart Game. Column (7)-(8) restrict attention to the decisions in the Restart Game of Group Adviser and Group GPT, respectively. *** p<0.01, ** p<0.05, * p<0.1.

Table B.9. Summary Statistics on the Coefficients of Contribution (Left) and Cognitive Uncertainty (Right)

| | N | Mean | SD | Min | Max | N | Mean | SD | Min | Max |
|---|---|---|---|---|---|---|---|---|---|---|
| **Pooled, All rounds, Total Subjects = 140, cognitive uncertainty makes participants:** | | | | | | | | | | |
| contribute more | 40 | 0.726 | 0.451 | 0.0700 | 2.370 | 40 | 26.27 | 15.97 | 1 | 68.50 |
| contribute less | 5 | -0.824 | 0.426 | -1.470 | -0.390 | 5 | 48.10 | 28.19 | 6 | 79.50 |
| **Pooled, First game, Total Subjects = 140, cognitive uncertainty makes participants:** | | | | | | | | | | |
| contribute more | 25 | 0.922 | 0.468 | 0.260 | 2.250 | 25 | 24.14 | 16.66 | 1 | 68.50 |
| contribute less | 11 | -1.056 | 0.375 | -1.680 | -0.610 | 11 | 41.45 | 19.68 | 6 | 79.50 |
| **Pooled, Restart game, Total Subjects = 140, cognitive uncertainty makes participants:** | | | | | | | | | | |
| contribute more | 27 | 0.895 | 0.712 | 0.250 | 3.890 | 27 | 23.72 | 14.79 | 5 | 67.50 |
| contribute less | 5 | -0.756 | 0.692 | -1.800 | -0.0700 | 5 | 53.30 | 23.35 | 19.50 | 79.50 |
| **Treatment Baseline, All rounds, Total Subjects = 44, cognitive uncertainty makes participants:** | | | | | | | | | | |
| contribute more | 15 | 0.760 | 0.529 | 0.320 | 2.370 | 15 | 26.57 | 16.65 | 1 | 67.50 |
| contribute less | 2 | -0.575 | 0.262 | -0.760 | -0.390 | 2 | 25.75 | 27.93 | 6 | 45.50 |
| **Treatment Baseline, First game, Total Subjects = 44, cognitive uncertainty makes participants:** | | | | | | | | | | |
| contribute more | 9 | 0.872 | 0.601 | 0.260 | 2.250 | 9 | 22.61 | 14.80 | 1 | 43.50 |
| contribute less | 3 | -1.063 | 0.413 | -1.540 | -0.810 | 3 | 30.33 | 21.29 | 6 | 45.50 |
| **Treatment Baseline, Restart game, Total Subjects = 44, cognitive uncertainty makes participants:** | | | | | | | | | | |
| contribute more | 9 | 0.763 | 0.452 | 0.250 | 1.450 | 0 | | | | |
| contribute less | 9 | 27.78 | 16.88 | 9.500 | 67.50 | 0 | | | | |
| **Treatment Adviser, Restart game, Total Subjects = 48, cognitive uncertainty makes participants:** | | | | | | | | | | |
| contribute more | 8 | 0.774 | 0.284 | 0.330 | 1.360 | 8 | 25.75 | 13.64 | 11.50 | 48.50 |
| contribute less | 2 | -0.840 | 0.255 | -1.020 | -0.660 | 2 | 60.75 | 26.52 | 42 | 79.50 |
| **Treatment GPT, Restart game, Total Subjects = 48, cognitive uncertainty makes participants:** | | | | | | | | | | |
| contribute more | 10 | 1.111 | 1.066 | 0.330 | 3.890 | 10 | 18.45 | 13.52 | 5 | 49 |
| contribute less | 3 | -0.700 | 0.956 | -1.800 | -0.0700 | 3 | 48.33 | 25.42 | 19.50 | 67.50 |

## Table B.10. Treatment Effect on Contribution

| GPT | Pooled | | | Inexperienced | | | Experienced | | |
|---|---|---|---|---|---|---|---|---|---|
| Treatment | GPT | Adviser | GPT | GPT | Adviser | GPT | GPT | Adviser | GPT |
| Control | Baseline | Baseline | Adviser | Baseline | Baseline | Adviser | Baseline | Baseline | Adviser |
| Sample in Restart Game | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| **DID (Round<=1)** | 9.566* | **9.103**** | 0.176 | 3.858 | 9.944 | -5.790 | **13.23**** | 8.814 | 3.888 |
| | (4.834) | **(4.236)** | (4.556) | (6.980) | (6.308) | (6.841) | **(6.586)** | (5.709) | (6.027) |
| *N* | 918 | **918** | 960 | 400 | 370 | 370 | **518** | 548 | 590 |
| Parallel test p value | 0.637 | **0.762** | 0.787 | 0.724 | 0.869 | 0.653 | **0.363** | 0.759 | 0.495 |
| **DID (Round<=2)** | 7.786* | **7.582**** | -0.198 | 6.095 | 7.753 | -1.455 | 8.631* | 7.429 | 0.572 |
| | (4.113) | **(3.645)** | (3.439) | (7.007) | (5.378) | (6.235) | (5.030) | (4.954) | (3.900) |
| *N* | 1,010 | **1,010** | 1,056 | 440 | 407 | 407 | 570 | 603 | 649 |
| Parallel test p value | 0.638 | **0.758** | 0.782 | 0.728 | 0.867 | 0.681 | 0.363 | 0.755 | 0.496 |
| **DID (Round<=3)** | 5.712 | 6.069* | -0.581 | 2.920 | 4.428 | -1.385 | 7.788* | 7.491* | -0.128 |
| | (3.762) | (3.403) | (2.957) | (6.885) | (5.465) | (5.609) | (4.110) | (4.255) | (3.149) |
| *N* | 1,102 | 1,102 | 1,152 | 480 | 444 | 444 | 622 | 658 | 708 |
| Parallel test p value | 0.627 | 0.758 | 0.773 | 0.759 | 0.869 | 0.690 | 0.363 | 0.756 | 0.494 |
| **DID (Round<=4)** | 4.912 | 5.228 | -0.453 | 2.740 | 4.427 | -1.619 | 6.545* | 6.059 | 0.201 |
| | (3.668) | (3.252) | (2.831) | (6.904) | (5.339) | (5.534) | (3.830) | (3.965) | (2.867) |
| *N* | 1,194 | 1,194 | 1,248 | 520 | 481 | 481 | 674 | 713 | 767 |
| Parallel test p value | 0.624 | 0.759 | 0.781 | 0.773 | 0.871 | 0.692 | 0.364 | 0.754 | 0.497 |
| **DID (Round<=5)** | 4.673 | 4.833 | -0.319 | 2.957 | 3.695 | -0.703 | 6.007 | 5.908 | -0.187 |
| | (3.448) | (3.201) | (2.567) | (6.309) | (5.188) | (4.648) | (3.764) | (3.926) | (2.882) |
| *N* | 1,286 | 1,286 | 1,344 | 560 | 518 | 518 | 726 | 768 | 826 |
| Parallel test p value | 0.625 | 0.761 | 0.782 | 0.773 | 0.872 | 0.682 | 0.364 | 0.756 | 0.496 |

Note. Average treatment effect on the treated (ATET) from observational data by difference in difference for panel data, that additionally control for the lagged average contribution from the group members, with robust standard error clustered (in parentheses) clustered at subject level. Estimate adjusted for covariates, panel effects, and time effects, and robust standard error cluster at subject level. Row parallel test p value reports the p value for parallel trends test (at round <=10) that the null hypothesis is linear trends are parallel. *** p<0.01, ** p<0.05, * p<0.1

Table B.11. Treatment Effect on Contribution, contd.

| GPT | Pooled | | | Inexperienced | | | Experienced | | |
|---|---|---|---|---|---|---|---|---|---|
| Treatment | GPT | Adviser | GPT | GPT | Adviser | GPT | GPT | Adviser | GPT |
| Control | Baseline | Baseline | Adviser | Baseline | Baseline | Adviser | Baseline | Baseline | Adviser |
| Sample in Restart Game | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| **DID (Round<=6)** | 3.975 | 4.266 | -0.438 | 2.637 | 3.148 | -0.497 | 5.017 | 5.328 | -0.527 |
| | (3.329) | (3.132) | (2.368) | (5.995) | (5.092) | (4.096) | (3.708) | (3.833) | (2.781) |
| *N* | 1,378 | 1,377 | 1,439 | 600 | 555 | 555 | 778 | 822 | 884 |
| Parallel test p value | 0.625 | 0.760 | 0.789 | 0.773 | 0.874 | 0.679 | 0.364 | 0.753 | 0.498 |
| **DID (Round<=7)** | 2.843 | 3.458 | -0.657 | 1.303 | 2.630 | -1.242 | 4.032 | 4.381 | -0.431 |
| | (3.188) | (3.022) | (2.276) | (5.620) | (4.838) | (3.829) | (3.670) | (3.762) | (2.746) |
| *N* | 1,470 | 1,468 | 1,534 | 640 | 592 | 592 | 830 | 876 | 942 |
| Parallel test p value | 0.626 | 0.753 | 0.803 | 0.770 | 0.871 | 0.662 | 0.364 | 0.746 | 0.502 |
| **DID (Round<=8)** | 2.441 | 3.295 | -0.858 | 0.657 | 1.905 | -1.167 | 3.883 | 4.617 | -0.764 |
| | (3.197) | (2.969) | (2.339) | (5.818) | (4.849) | (4.104) | (3.513) | (3.611) | (2.715) |
| *N* | 1,562 | 1,560 | 1,630 | 680 | 629 | 629 | 882 | 931 | 1,001 |
| Parallel test p value | 0.624 | 0.753 | 0.804 | 0.777 | 0.871 | 0.667 | 0.365 | 0.745 | 0.503 |
| **DID (Round<=9)** | 2.278 | 3.290 | -1.021 | 0.257 | 1.872 | -1.504 | 3.854 | 4.639 | -0.807 |
| | (3.168) | (2.958) | (2.288) | (5.767) | (4.963) | (4.008) | (3.432) | (3.484) | (2.677) |
| *N* | 1,654 | 1,652 | 1,726 | 720 | 666 | 666 | 934 | 986 | 1,060 |
| Parallel test p value | 0.624 | 0.754 | 0.802 | 0.773 | 0.872 | 0.680 | 0.365 | 0.745 | 0.504 |
| **DID (Round<=10)** | 2.291 | 3.403 | -1.119 | -0.0983 | 2.060 | -2.016 | 4.169 | 4.711 | -0.605 |
| | (3.134) | (2.938) | (2.282) | (5.758) | (5.039) | (3.943) | (3.328) | (3.364) | (2.712) |
| *N* | 1,746 | 1,744 | 1,822 | 760 | 703 | 703 | 986 | 1,041 | 1,119 |
| Parallel test p value | 0.625 | 0.755 | 0.798 | 0.779 | 0.873 | 0.688 | 0.365 | 0.746 | 0.503 |

Note. Average treatment effect on the treated (ATET) from observational data by difference in difference for panel data, that additionally control for the lagged average contribution from the group members, with robust standard error clustered (in parentheses) clustered at subject level. Estimate adjusted for covariates, panel effects, and time effects, and robust standard error cluster at subject level. Row parallel test p value reports the p value for parallel trends test (at round <=10) that the null hypothesis is linear trends are parallel. *** p<0.01, ** p<0.05, * p<0.1

Table B.12. Treatment Effect on Cognitive Uncertainty

| GPT | Pooled | | | Inexperienced | | | Experienced | | |
|---|---|---|---|---|---|---|---|---|---|
| Treatment | GPT | Adviser | GPT | GPT | Adviser | GPT | GPT | Adviser | GPT |
| Control | Baseline | Baseline | Adviser | Baseline | Baseline | Adviser | Baseline | Baseline | Adviser |
| Sample in Restart Game | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| **DID (Round<=1)** | 6.777* | **8.125**** | -1.338 | **10.87**** | 5.549 | 5.114 | 3.952 | 9.441 | -5.657 |
| | (3.540) | **(3.754)** | (3.514) | **(5.170)** | (4.009) | (5.440) | (5.194) | (5.898) | (4.341) |
| *N* | 918 | **918** | 960 | **400** | 370 | 370 | 518 | 548 | 590 |
| Parallel test p value | 0.652 | **0.934** | 0.601 | **0.616** | 0.417 | 0.608 | 0.857 | 0.466 | 0.297 |
| **DID (Round<=2)** | 4.817* | 4.594* | 0.177 | 6.667 | 1.278 | 5.386 | 3.681 | 6.582 | -3.170 |
| | (2.824) | (2.678) | (2.860) | (4.684) | (2.983) | (4.684) | (3.742) | (4.034) | (3.482) |
| *N* | 1,010 | 1,010 | 1,056 | 440 | 407 | 407 | 570 | 603 | 649 |
| Parallel test p value | 0.651 | 0.935 | 0.594 | 0.616 | 0.416 | 0.635 | 0.856 | 0.466 | 0.297 |
| **DID (Round<=3)** | 3.515 | 3.987 | -0.508 | 1.388 | 0.106 | 1.331 | 5.570* | 7.001* | -1.747 |
| | (2.514) | (2.428) | (2.417) | (4.011) | (3.023) | (4.124) | (3.319) | (3.507) | (2.928) |
| *N* | 1,102 | 1,102 | 1,152 | 480 | 444 | 444 | 622 | 658 | 708 |
| Parallel test p value | 0.644 | 0.930 | 0.583 | 0.591 | 0.420 | 0.654 | 0.858 | 0.461 | 0.295 |
| **DID (Round<=4)** | 4.549* | 4.075* | 0.396 | 2.771 | 2.737 | -0.0322 | 6.234* | 5.467* | 0.521 |
| | (2.388) | (2.259) | (2.331) | (3.850) | (3.109) | (4.235) | (3.185) | (3.230) | (2.703) |
| *N* | 1,194 | 1,194 | 1,248 | 520 | 481 | 481 | 674 | 713 | 767 |
| Parallel test p value | 0.648 | 0.929 | 0.586 | 0.604 | 0.426 | 0.661 | 0.859 | 0.464 | 0.295 |
| **DID (Round<=5)** | 4.269* | **4.359**** | -0.178 | 1.495 | 3.330 | -1.858 | **6.701**** | 5.635* | 0.747 |
| | (2.263) | **(2.190)** | (2.255) | (3.618) | (3.254) | (4.166) | **(3.001)** | (2.992) | (2.599) |
| *N* | 1,286 | **1,286** | 1,344 | 560 | 518 | 518 | **726** | 768 | 826 |
| Parallel test p value | 0.649 | **0.929** | 0.583 | 0.605 | 0.424 | 0.666 | **0.859** | 0.462 | 0.296 |

Note. Average treatment effect on the treated (ATET) from observational data by difference in difference for panel data, that additionally control for the lagged average contribution from the group members, with robust standard error clustered (in parentheses) clustered at subject level. Estimate adjusted for covariates, panel effects, and time effects, and robust standard error cluster at subject level. Row parallel test p value reports the p value for parallel trends test (at round <=10) that the null hypothesis is linear trends are parallel. *** p<0.01, ** p<0.05, * p<0.1

Table B.13. Treatment Effect on Cognitive Uncertainty, contd.

| GPT | Pooled | | | Inexperienced | | | Experienced | | |
|---|---|---|---|---|---|---|---|---|---|
| Treatment | GPT | Adviser | GPT | GPT | Adviser | GPT | GPT | Adviser | GPT |
| Control | Baseline | Baseline | Adviser | Baseline | Baseline | Adviser | Baseline | Baseline | Adviser |
| Sample in Restart Game | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| **DID (Round<=6)** | 3.649 | 3.895* | -0.292 | 0.157 | 3.385 | -3.199 | **6.639**** | 4.974* | 1.363 |
| | (2.204) | (2.255) | (2.210) | (3.356) | (3.619) | (4.270) | **(3.010)** | (2.954) | (2.460) |
| *N* | 1,378 | 1,377 | 1,439 | 600 | 555 | 555 | **778** | 822 | 884 |
| Parallel test p value | 0.650 | 0.933 | 0.587 | 0.610 | 0.419 | 0.656 | **0.859** | 0.463 | 0.297 |
| **DID (Round<=7)** | 3.039 | 3.234 | -0.250 | -0.664 | 3.426 | -4.052 | 6.160* | 3.891 | 1.924 |
| | (2.235) | (2.336) | (2.152) | (3.217) | (3.633) | (4.125) | (3.160) | (3.135) | (2.404) |
| *N* | 1,470 | 1,468 | 1,534 | 640 | 592 | 592 | 830 | 876 | 942 |
| Parallel test p value | 0.650 | 0.932 | 0.583 | 0.615 | 0.415 | 0.640 | 0.858 | 0.459 | 0.295 |
| **DID (Round<=8)** | 2.141 | 2.367 | -0.270 | -1.789 | 3.385 | -5.145 | 5.375* | 2.535 | 2.553 |
| | (2.246) | (2.332) | (2.099) | (3.175) | (3.585) | (3.965) | (3.196) | (3.131) | (2.339) |
| *N* | 1,562 | 1,560 | 1,630 | 680 | 629 | 629 | 882 | 931 | 1,001 |
| Parallel test p value | 0.648 | 0.931 | 0.582 | 0.612 | 0.415 | 0.638 | 0.859 | 0.457 | 0.294 |
| **DID (Round<=9)** | 2.121 | 1.861 | 0.185 | -2.122 | 3.337 | -5.507 | 5.634* | 1.869 | 3.491 |
| | (2.353) | (2.403) | (2.154) | (3.391) | (3.612) | (4.025) | (3.312) | (3.233) | (2.380) |
| *N* | 1,654 | 1,652 | 1,726 | 720 | 666 | 666 | 934 | 986 | 1,060 |
| Parallel test p value | 0.641 | 0.924 | 0.577 | 0.594 | 0.424 | 0.658 | 0.861 | 0.455 | 0.293 |
| **DID (Round<=10)** | 1.506 | 1.604 | -0.185 | -3.619 | 2.603 | -6.272 | 5.664* | 1.966 | 3.422 |
| | (2.415) | (2.481) | (2.206) | (3.535) | (3.775) | (4.103) | (3.324) | (3.290) | (2.449) |
| *N* | 1,746 | 1,744 | 1,822 | 760 | 703 | 703 | 986 | 1,041 | 1,119 |
| Parallel test p value | 0.641 | 0.922 | 0.574 | 0.600 | 0.422 | 0.651 | 0.862 | 0.452 | 0.292 |

Note. Average treatment effect on the treated (ATET) from observational data by difference in difference for panel data, that additionally control for the lagged average contribution from the group members, with robust standard error clustered (in parentheses) clustered at subject level. Estimate adjusted for covariates, panel effects, and time effects, and robust standard error cluster at subject level. Row parallel test p value reports the p value for parallel trends test (at round <=10) that the null hypothesis is linear trends are parallel. *** p<0.01, ** p<0.05, * p<0.1

## Appendix C Instruction

The following pages show the screenshot for the experiment. For those without bracket, it means that the page is showed to all treatment.

# Instruction

There are 10 rounds of games.

In each round, you will be randomly assigned to interact in a group with 3 of the other participants.

There are two tasks in each round.

## Task 1

Each of the 4 participants in your group will receive an endowment of 20 points.

Every one of you must decide how much of your endowment of 20 points to contribute to the "group project". The contribution must be an integer between 0 to 20 (including 0 and 20).

All points contributed to the "group project" are multiplied by 1.6 and then split evenly among the 4 group members.

Therefore, in each round, your payoff will be the sum of the part of your 20 points endowment you keep for yourself, plus your share of the group project, i.e.,

$$\boxed{\begin{array}{c}\text{Your Payoff in Points} \\ = (20-\text{your contribution to group project}) + 0.4\times(\text{total contributions to the group project})\end{array}}$$

## Task 2

Next, you will indicate how certain are you that you would actually want to contribute to the group project at the amount you indicate.

| 0% | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% | 100% |
|---|---|---|---|---|---|---|---|---|---|---|
| very uncertain | | | | | | | | | | completely certain |

## Result

After all members have made their decisions, you will be able to view the information about your payoff from the current round.

## Random Grouping in Each Round

Then, the next round will begin, with the composition of the groups changing for each round, which means that you will be playing with different participants in every round.

## Payoff

At the end of the experiment, the sum of your 10 payoffs will be converted to Singapore dollars (SGD) at a rate of 1 Point = 0.025 SGD, which is your final earning for the experiment.

You will be asked some demographic questions at the end of the experiment. Then, you will be paid in a private manner.

Next

## Control Question 1/5

1. Each of the 4 group members has an endowment of 20 points. Suppose nobody (including you) contributes any points to the group project.

What is your payoff (in points)?

What is the payoff of each of the other group members (in points)?

*Hints*:

Your Payoff
= (20−your contribution to group project) + 0.4×(total contributions to the group project)

Next

## Control Question 2/5

2. Each of the 4 group members has an endowment of 20 points. Suppose you contribute 20 points to the group project. All other 3 group members each contribute 20 points to the group project.

What is your payoff (in points)?

What is the payoff of each of the other group members (in points)?

*Hints*:

Your Payoff
= (20−your contribution to group project) + 0.4×(total contributions to the group project)

Next

# Control Question 3/5

3. Each of the 4 group members has an endowment of 20 points. Suppose the other 3 group members contribute a total of 30 points to the group project.

What is your payoff (in points) if you do not contribute any points to the group project?

What is your payoff (in points) if you contribute 15 points to the group project?

*Hints*:

Your Payoff
= (20−your contribution to group project) + 0.4×(total contributions to the group project)

Next

# Control Question 4/5

4. Each of the 4 group members has an endowment of 20 points. Suppose you contribute 8 points to the group project.

What is your payoff (in points) if all other group members together contribute a total of 7 points to the group project?

What is your payoff (in points) if all other group members together contribute a total of 22 points to the group project?

*Hints*:

Your Payoff
= (20−your contribution to group project) + 0.4×(total contributions to the group project)

Next

## Control Question 5/5

5.Suppose you are 80% certain that you would actually want to contribute to the group project the amount you indicate. Which button should you click in this case?

| 0% | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% | 100% |
|---|---|---|---|---|---|---|---|---|---|---|

**very uncertain** **completely certain**

Next

## Round 1 of 10

### Task 1

| | |
|---|---|
| Your Endowment in Round 1 (in Points): | 20 |
| Your Contribution to the Group Project in Round 1 (in Points): | |

Next

## Round 1 of 10

### Task 2

Your decision on the previous screen indicate that you would like to contribute **5 point(s)** to the group project.

**How certain** are you that you would actually want to contribute somewhere **between 4 point(s) and 6 point(s)** to the project?

| 0% | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% | 100% |
|---|---|---|---|---|---|---|---|---|---|---|

**very uncertain** **completely certain**

Next

# Round 1 of 10

Time left to complete this page: **0:07**

## Payoff Summary

| | |
|---|---|
| Your contribution to the group project: | 5 points |
| Total contribution to the group project from your group: | 39 points |
| The amount of points you keep for yourself: | 15 points |
| Payoff from the group project: | 15.60 points |
| Your total payoff in Round 1 (round to integer): | 31 points |

Next

[Group Baseline]

# Instruction

The second part of the experiment consists of 10 rounds of games. The instructions for this part are exactly the same as the ones for the 10 rounds in the previous part. You can revise it in the "Recap of Instruction in Part I" Section.

## Recap of Instruction in Part I

In each round, you will be randomly assigned to interact in a group with 3 of the other participants.

There are two tasks in each round.

### Task 1

Each of the 4 participants in your group will receive an endowment of 20 points.

Every one of you must decide how much of your endowment of 20 points to contribute to the "group project". The contribution must be an integer between 0 to 20 (including 0 and 20).

All points contributed to the "group project" are multiplied by 1.6 and then split evenly among the 4 group members.

Therefore, in each round, your payoff will be the sum of the part of your 20 points endowment you keep for yourself, plus your share of the group project, i.e.,

Your Payoff in Points
= (20−your contribution to group project) + 0.4×(total contributions to the group project)

### Task 2

Next, you will indicate how certain are you that you would actually want to contribute to the group project at the amount you indicate.

| 0% | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% | 100% |
|---|---|---|---|---|---|---|---|---|---|---|
| very uncertain | | | | | | | | | | completely certain |

### Result

After all members have made their decisions, you will be able to view the information about your payoff from the current round.

### Random Grouping in Each Round

Then, the next round will begin, with the composition of the groups changing for each round, which means that you will be playing with different participants in every round.

### Payoff

At the end of the experiment, the sum of your 10 payoffs will be converted to Singapore dollars (SGD) at a rate of 1 Point = 0.025 SGD, which is your final earning for the experiment.

You will be asked some demographic questions at the end of the experiment. Then, you will be paid in a private manner.

Next

[Group Adviser]

# Instruction

The second part of the experiment consists of 10 rounds of games. The instructions for this part are exactly the same as the ones for the 10 rounds in the previous part. You can revise it in the "Recap of Instruction in Part I" Section.

## Recommendation by the Adviser

The only difference is that, prior to the beginning of this part, we provide you with the contribution decisions made by the adviser.

The adviser was given the same instructions as you in the last part of the experiment and was instructed to make the decisions.

**The contribution level chosen by the adviser is usually around 13 out of 20 points of the endowment.**

## Recap of Instruction in Part I

In each round, you will be randomly assigned to interact in a group with 3 of the other participants.

There are two tasks in each round.

### Task 1

Each of the 4 participants in your group will receive an endowment of 20 points.

Every one of you must decide how much of your endowment of 20 points to contribute to the "group project". The contribution must be an integer between 0 to 20 (including 0 and 20).

All points contributed to the "group project" are multiplied by 1.6 and then split evenly among the 4 group members.

Therefore, in each round, your payoff will be the sum of the part of your 20 points endowment you keep for yourself, plus your share of the group project, i.e.,

Your Payoff in Points
= (20−your contribution to group project) + 0.4×(total contributions to the group project)

### Task 2

Next, you will indicate how certain are you that you would actually want to contribute to the group project at the amount you indicate.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 0% | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% | 100% |
| very uncertain | | | | | | | | | | completely certain |

### Result

After all members have made their decisions, you will be able to view the information about your payoff from the current round.

### Random Grouping in Each Round

Then, the next round will begin, with the composition of the groups changing for each round, which means that you will be playing with different participants in every round.

### Payoff

At the end of the experiment, the sum of your 10 payoffs will be converted to Singapore dollars (SGD) at a rate of 1 Point = 0.025 SGD, which is your final earning for the experiment.

You will be asked some demographic questions at the end of the experiment. Then, you will be paid in a private manner.

Next

[Group GPT, link = https://en.wikipedia.org/wiki/GPT-3]

# Instruction

The second part of the experiment consists of 10 rounds of games. The instructions for this part are exactly the same as the ones for the 10 rounds in the previous part. You can revise it in the "Recap of Instruction in Part I" Section.

## Recommendation by GPT-3.5

The only difference is that, prior to the beginning of this part, we provide you with the contribution decisions made by Generative Pre-trained Transformer 3.5 (GPT-3.5).

GPT-3.5 is a subclass of GPT-3 Models created by OpenAI that implements AI and focuses on natural language understanding and generation. You can read about it on Wikipedia here.

GPT-3.5 was given the same instructions as you in the last part of the experiment and was instructed to make the decisions.

**The contribution level chosen by GPT-3.5 is usually around 13 out of 20 points of the endowment.**

## Recap of Instruction in Part I

In each round, you will be randomly assigned to interact in a group with 3 of the other participants.

There are two tasks in each round.

### Task 1

Each of the 4 participants in your group will receive an endowment of 20 points.

Every one of you must decide how much of your endowment of 20 points to contribute to the "group project". The contribution must be an integer between 0 to 20 (including 0 and 20).

All points contributed to the "group project" are multiplied by 1.6 and then split evenly among the 4 group members.

Therefore, in each round, your payoff will be the sum of the part of your 20 points endowment you keep for yourself, plus your share of the group project, i.e.,

Your Payoff in Points
= (20−your contribution to group project) + 0.4×(total contributions to the group project)

### Task 2

Next, you will indicate how certain are you that you would actually want to contribute to the group project at the amount you indicate.

| 0% | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% | 100% |
|---|---|---|---|---|---|---|---|---|---|---|
| very uncertain | | | | | | | | | | completely certain |

### Result

After all members have made their decisions, you will be able to view the information about your payoff from the current round.

### Random Grouping in Each Round

Then, the next round will begin, with the composition of the groups changing for each round, which means that you will be playing with different participants in every round.

### Payoff

At the end of the experiment, the sum of your 10 payoffs will be converted to Singapore dollars (SGD) at a rate of 1 Point = 0.025 SGD, which is your final earning for the experiment.

You will be asked some demographic questions at the end of the experiment. Then, you will be paid in a private manner.

Next

# Round 1 of 10

## Task 2

Your decision on the previous screen indicate that you would like to contribute **5 point(s)** to the group project.

**How certain** are you that you would actually want to contribute somewhere **between 4 point(s) and 6 point(s)** to the project?

| 0% | 10% | 20% | 30% | 40% | 50% | 60% | 70% | 80% | 90% | 100% |
|---|---|---|---|---|---|---|---|---|---|---|

**very uncertain** **completely certain**

Next

# Round 1 of 10

Time left to complete this page: **0:07**

## Payoff Summary

| | |
|---|---|
| Your contribution to the group project: | 5 points |
| Total contribution to the group project from your group: | 39 points |
| The amount of points you keep for yourself: | 15 points |
| Payoff from the group project: | 15.60 points |
| Your total payoff in Round 1 (round to integer): | 31 points |

Next

## Appendix D GPT Prompt in Public goods game

The following shows the prompt we give to GPT on 14 June 2023, using GPT-3.5-turbo model.

**System role:** "I want you to act as a human decision maker. You will be given 10 rounds of decision-making tasks, and will be responsible for making decisions. You should use your best judgement to come up with a solution that you like the most. My first request is: You should provide your answer in every round. If you do not provide an answer, I will assume that you make a random choice. This prompt serves as the background information for the user prompt later. No answer from this prompt is required."

**Assistant role:** "In every round, the decision maker is randomly matched with 3 other new subjects to form a group of 4 and there is no feedback across rounds, as feedback is trivial. Every member of the group needs to decide how many of the 20 points he/she wants to contribute to the group project. The contribution to the group project must be an integer between 0 and 20 (including 0 and 20). All contributions to the group project will be multiplied by 1.6, and split evenly among the 4 group members, yielding the return from the group project for each member. In every round, each member's payoff will be the sum of the part of the 20 points endowment he/she keeps for himself/herself (i.e., did not contribute to the group project), plus the return he/she gets from the group project. Please do not assume other group members will make the same contribution to the group project as you do. This prompt serves as the background information for the user prompt later. No answer from this prompt is required."

**User role:** In Round $N \in [1, 10]$, How many points would you like to contribute to the group project? How many points you keep for yourself? Please input the answers in integer between 0 to 20 (include 0 and 20).

Table D.1. Descriptive statistics

| | N | Mean | SD | Range |
|---|---|---|---|---|
| All sample | 150 | 12.64 | 4.32 | 0–20 |
| Temperature = 0 | 50 | 11.26 | 4.55 | 0–20 |
| Temperature = 0.5 | 50 | 13.72 | 4.38 | 8–20 |
| Temperature = 1 | 50 | 12.94 | 3.69 | 8–20 |

# Appendix E Complexity Experiments

## Appendix E.1 Experimental Design

The main study focuses on the correlational relationship between cognitive uncertainty and contribution behavior. In this causal study, we implement additional treatment arms in this section to manipulate the complexity of the task by displaying the efficiency factor in a mathematical equation. The complexity experiments are conducted following a similar fashion as in Enke et al. (2023). Specifically, we have made several changes to the design of the complexity experiments.

First, we no longer conduct the Treatment Baseline in the first game, nor do we include a surprise restart; instead, we impose the treatment from the beginning. This change is due to the first 10 rounds potentially lowering participants' cognitive uncertainty as they learn. Therefore, even if the causal implementation is successful, we may still not be able to observe a significant increase in cognitive uncertainty. Instead, we now have all groups in the Complexity Experiments engage in only 10 rounds of the game.

Secondly, we vary the multiplier in each round and control its effect in our regression. An alternative approach would be to apply different equations for the same multiplier, say M = 1.6. While this allows us to make comparisons with our treatment in the main design, it can make the setting feel unnatural, and, more importantly, it exposes us to the risk of annoying the subjects because they might feel as though they have been misled once they discover the truth.

We design two treatment groups in the Complexity Experiment. The screenshot of the interface can be found in Figure E.1.

In the Treatment Complexity Number, we vary the multiplier over 10 rounds. The multiplier does not differ significantly from that in the Treatment Baseline, where it ranges between 1.1 and 2.0, or, equivalently, a maximum difference of 0.225 in MPCR. The specific configuration of the multiplier can be found in Table E.1.

In the Treatment Complexity Equation, we incorporate the multiplier into a complex mathematical equation and impose a 25-second time limit on decision-making[12]. This setup aligns with the complexity treatment described in Enke et al. (2023), which aims to prevent participants from spending significantly more time on task completion, potentially resulting in the absence of a significant effect on cognitive uncertainty. We formulate the hypothesis in Experiment Complexity as follows:

**Hypothesis E1:** *A manipulation on the complexity of the multiplier increases cognitive uncertainty.*
**Hypothesis E2:** *A manipulation on the complexity of the multiplier increases contribution.*

[12]The management for when timeout happens follows all the other treatment in this paper, as discussed in Table B.2

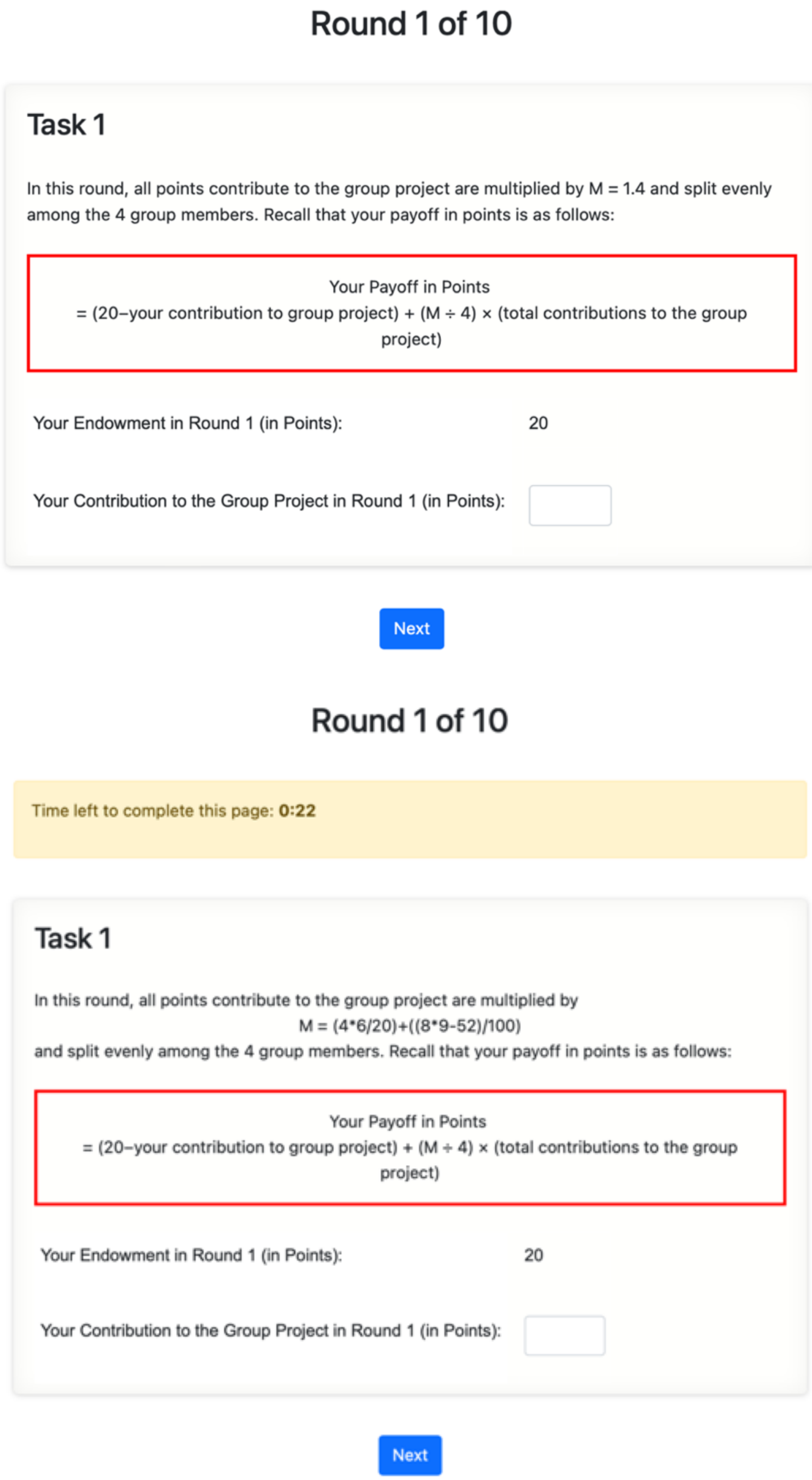



Figure E.1. Screenshot of an example decision screen in Complexity Number (upper panel) and Complexity Equation (bottom panel).

Table E.1. Equations with Round and Multiplier

| Round | Multiplier | Equation |
|---|---|---|
| 1 | 1.4 | $4*6/20+8*9-52/100$ |
| 2 | 1.2 | $0.6*3/2+5*0.08-0.1$ |
| 3 | 1.5 | $0.5*6/5+2*0.4+0.1$ |
| 4 | 1.7 | $30*5/100+8*0.5/40+0.1$ |
| 5 | 1.6 | $4*2-6-4*0.2+0.4$ |
| 6 | 1.9 | $3*0.9-1.0+5*6/100-0.1$ |
| 7 | 1.1 | $8*0.6/2-0.4*5/2-0.3$ |
| 8 | 2.0 | $0.8*3+0.6-6*0.5/2+0.5$ |
| 9 | 1.3 | $(30*5-20)/((2.4+7.6)*10)$ |
| 10 | 1.8 | $(20*9-90)/((6+4)*5)$ |

### Appendix E.2 Timeout Instances and Invalid Decisions in Complexity Equation

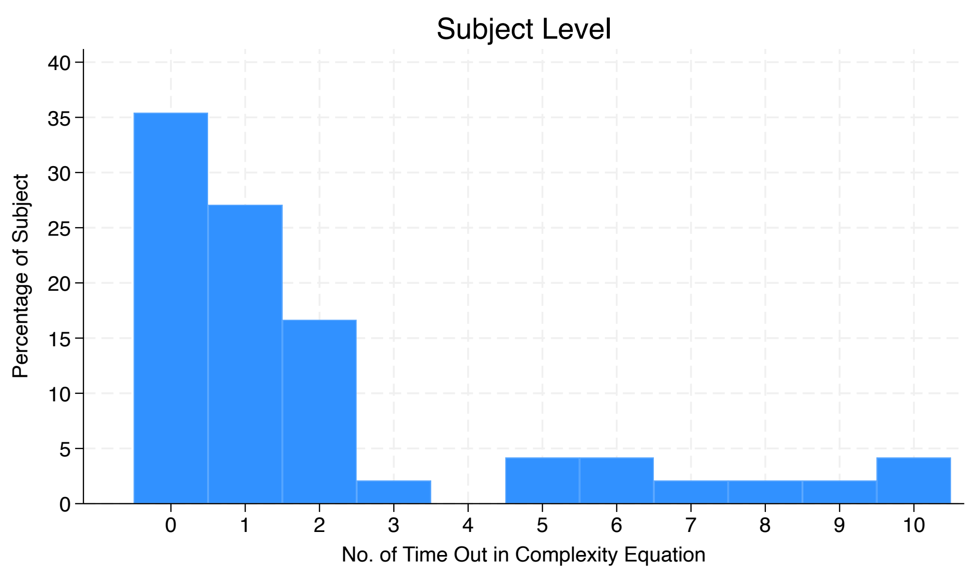



Figure E.2. Histogram of number of timeout instance in Complexity Equation on the subject level

There are 48 subjects participate in each of the treatment.

The first finding is that due to the hard time binding in Treatment Complexity Equation, there were 98 out of 480 instances where subjects failed to submit a contribution decision within the 25-second time limit in Complexity Equation.

As tabulated in Table E.2, each subject had approximately 2 instances out of 10 on average where they failed to submit a contribution in time in the Complexity Equation. As showed in Figure E.2, only 17 out of 48 subjects in the Complexity Equation managed to submit their contribution decisions on time for all the periods, and there are 2 subjects failing to submit every decision. By contrast, all subjects succeed to submit all contribution decisions in Complexity Number.

Table E.2. Summary Statistics on Timeout Instance

| | Observation | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|---|
| *No. of Timeout Instance* | | | | | |
| Complexity Number | 480 | 0 | 0 | 0 | 0 |
| Complexity Equation | 480 | 0.204 | 0.404 | 0 | 1 |
| *Average Timeout Instance Per Subject* | | | | | |
| Complexity Number | 48 | 0 | 0 | 0 | 0 |
| Complexity Equation | 48 | 2.042 | 2.805 | 0 | 10 |

Table E.3. Timeout, Completion Time, and Total Payoff across Treatments

| Dependent Variable | Timeout<br>(1) | Completion Time (in Min)<br>(2) | Total Payoff<br>(3) |
|---|---|---|---|
| Round | -0.01*** | | |
| | (0.00) | | |
| 1 if Complexity Equation | 0.21*** | -2.42*** | -33.78*** |
| | (0.04) | (0.38) | (11.45) |
| Multiplier | -0.01 | | |
| | (0.02) | | |
| Observations | 960 | 96 | 96 |
| R-Squared | 0.20 | 0.38 | 0.24 |

Note: OLS estimates with robust standard errors (in parentheses) are clustered at the subject level and OLS in column (1)-(2). Column (3)-(4) look at the difference in completion time and payoff when comparing the two treatments, controlling for the demographics variables including major, gender, age, whether received statistics training, experience with economics experiment, investment experiment, and use of GPT.. *** p<0.01, ** p<0.05, * p<0.1.

The observation is supported by Table E.3 Column (1)[13]. We find that the failure to submit decisions on time is significantly higher in the Complexity Equation treatment. Meanwhile, there is a decreasing pattern of time out incidence as the game repeats. Consequently, subjects in Complexity Equation earned less than their counterparts in the Complexity Number, and it would require them a total of 2.42 minutes more (equivalent to 14.52 seconds more for each decision) to submit all their decisions in every round.

### Appendix E.3 Non-binding Complexity Equation Soft

In response to the high incidence of missing values in the Complexity Equation treatment, we introduced an additional treatment called Complexity Equation Soft. In this treatment, subjects were provided with a non-binding time limit of 25 seconds to submit their decisions. An example screenshot can be found in Figure E.3. Subjects were still required to calculate the multiplier using a complex mathematical equation, similar to that in the Complexity treatment. However, in this case, they would not receive a payment of zero, and the page would not be automatically submitted, if they failed to input a contribution within 25 seconds. Instead, a pop-up window would notify them that their time had run out, and they needed to make a decision as soon as possible. A total of 40 subjects participate in the treatment.

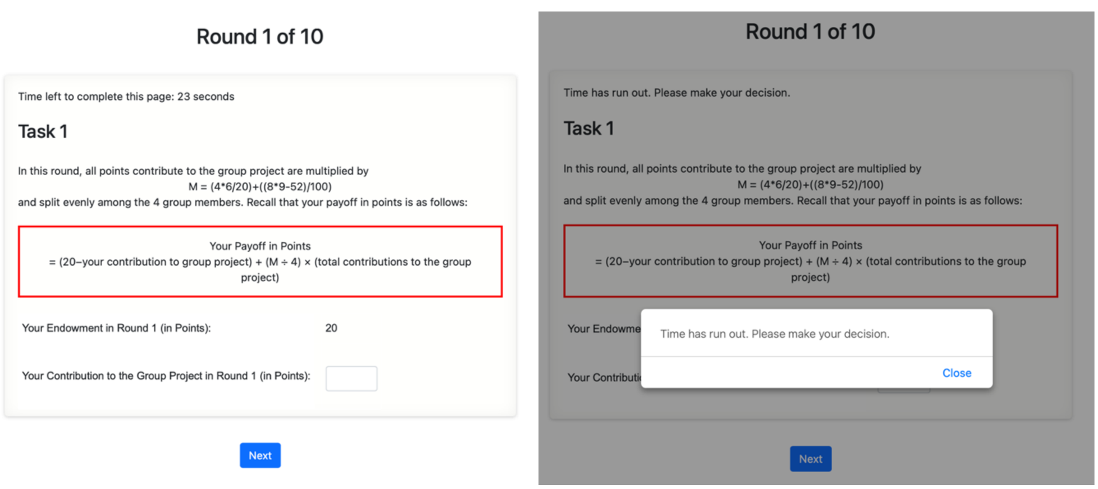



Figure E.3. Screenshot of an example decision screen in Complexity Number (upper panel) and Complexity Equation (bottom panel).

[13]It's important to note that we have excluded the control of lagged average contribution from other group members. This is because including the lagged payment control would require discarding observations from the first one periods, which is critical in this section when we vary multiplier in each round, because it would also entail discarding all observations with two unique multipliers.

### Appendix E.4 Experimental Results

We depict the average cognitive uncertainty and contribution across the treatment In Figure E.4. Both contribution and cognitive uncertainty is the highest in Complexity Equation Soft, and lowest in Complexity Number. Overall, we fail to find a significant difference between Complexity equation and Complexity Number on cognitive uncertainty ($t$-test: $t$=-0.1163, $p$=0.9074; Rank-sum: $z$=0.932, $p$=0.3513). However, we do find that the cognitive uncertainty is higher in Complexity Equation Soft than that of Complexity Number ($t$-test: $t$= 2.3737, $p$=0.0178; Rank-sum: $z$=2.009, $p$=0.0446). And the contribution in the both Complexity Equation ($t$-test: $t$=4.1164, $p$=0.0000; Rank-sum: $z$=3.400, $p$=0.0007) and Complexity Equation Soft (t-test: $t$=8.0275, $p$=0.0000; Rank-sum: $z$=8.329, $p$=0.0000) is significantly higher than that of Complexity Number.

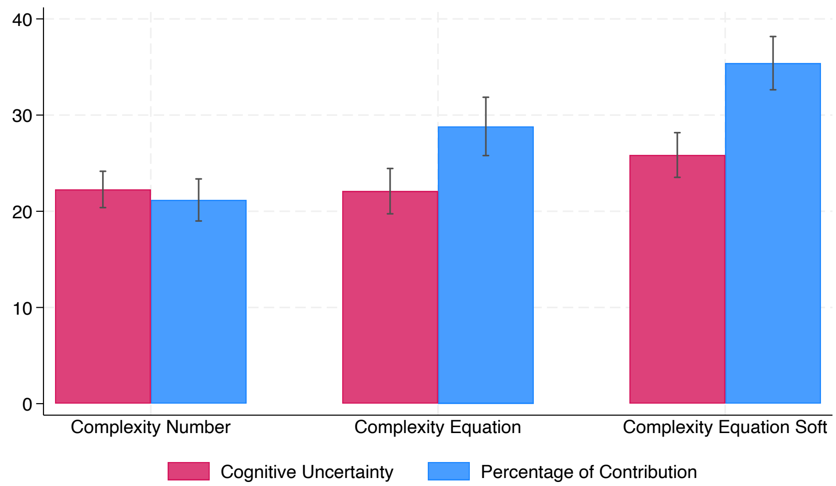



Figure E.4. Screenshot of an example decision screen in Complexity Number (upper panel) and Complexity Equation (bottom panel).

In summary, all the mean comparisons are consistent with our hypothesis, where complex mathematical function manipulation would increase both cognitive uncertainty and contribution; except for the comparison between Complexity Equation and Complexity Number regarding cognitive uncertainty. However, intuitively, a complex mathematical function manipulation with an additional binding time limit should lead to a larger cognitive uncertainty due to the consequence of zero payoff. Therefore, we conclude that the insignificant result is likely due to attrition in data collection when subjects become too confused about the task. Consequently, we failed to collect data on cognitive uncertainty and contribution when this confusion occurred.

**Result E.1:** *When manipulating the multiplier to be more complicated in a between-subject fashion, it leads subjects to be more cognitively uncertain about their contribution decision.*

Figure E.5 visualizes the contribution pattern, and Table E.4 summarizes the results. In support of our hypothesis, we find that a complexity manipulation increases contribution, as illustrated by the higher reddish dots in Figure E.5 compared to the blue dots. This observation aligns with the positive coefficient of the complexity manipulation as shown in the OLS estimates in Columns (1) and (2) of Table E.4. Similarly, the negative coefficient in Columns (5) and (6) indicates that contributions are closer to half of the endowment when the multiplier is in the form of a complicated function. However, the result is not statistically significant in the Complexity Equation treatment, probably due to are attritions in the contribution decisions as discussed.

Furthermore, the contribution is positively correlated with the multiplier, but the inelasticity pattern, as indicated by the coefficient of the interaction term, is not statistically significant at the 5% level. This is possibly due to our design, where the contribution in our configuration is smaller than half of the endowment even without cognitive uncertainty. In turn, cognitive uncertainty would only make the contribution decision shift closer to the default (i.e., increase), not decrease. Hence, we were unable to identify an inelastic pattern.

**Result E2:** *When manipulating the multiplier to be more complicated in a between-subject fashion, it leads subjects to contribute at a higher level that is closer to the cognitive default of half-of-endowment.*

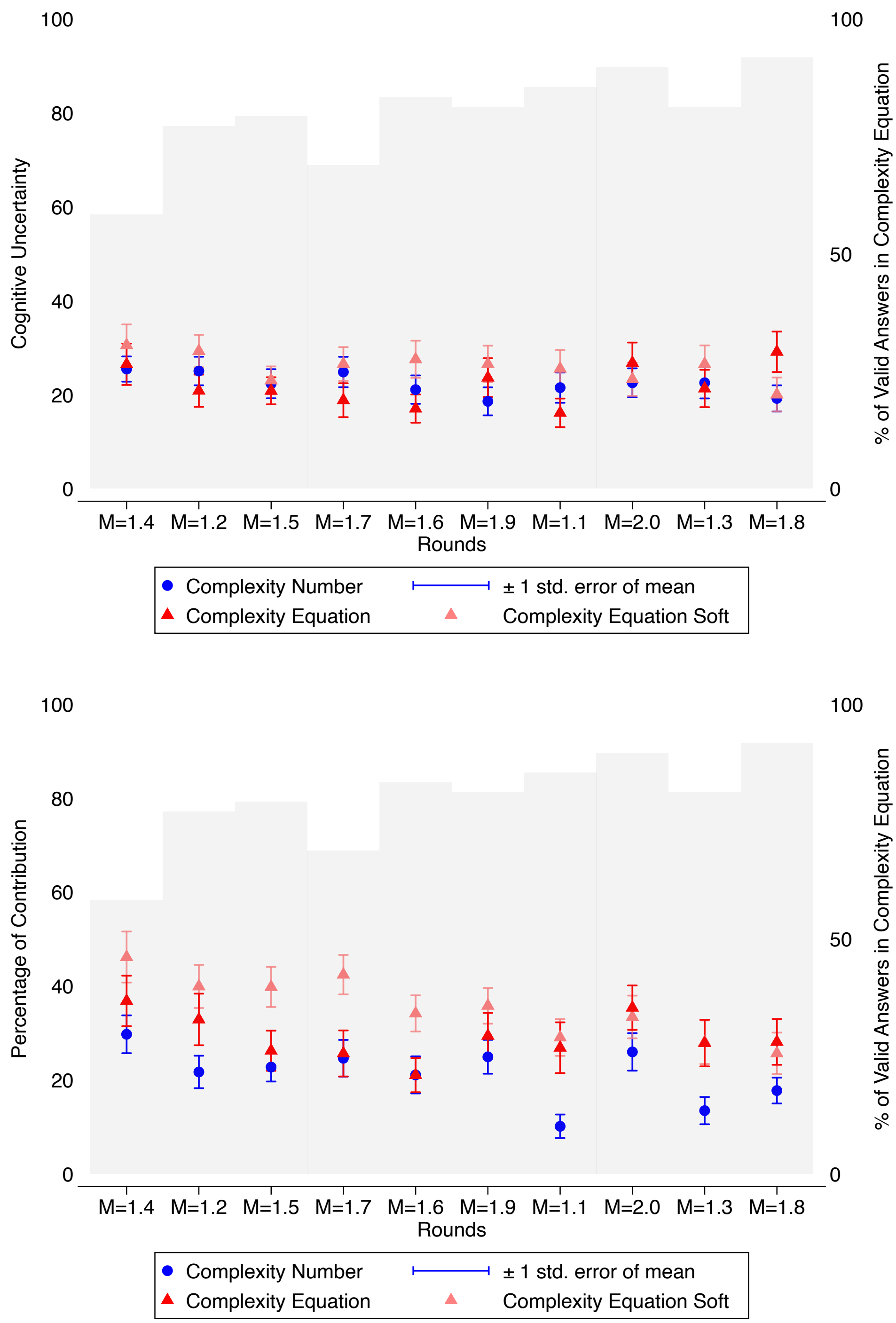



Figure E.5. Percentage of Contribution in Complexity Number (Number of Subjects = 48), Complexity Equation (Number of Subjects = 46, where 2 subject do not submit any contribution decision), and Complexity Equation Soft (Number of Subjects = 40). Percentage of contribution are computed using their contribution divided by endowment of 20 points. Whiskers show standard error bars.

Table E.4. Complexity Manipulations

| Dependent Variable | Timeout<br>(1) | Completion Time (in Min)<br>(2) | Total Payoff<br>(3) |
|---|---|---|---|
| Round | -0.01***<br>(0.00) | | |
| 1 if Complexity Equation | 0.21***<br>(0.04) | -2.42***<br>(0.38) | -33.78***<br>(11.45) |
| Multiplier | -0.01<br>(0.02) | | |
| Observations | 960 | 96 | 96 |
| R-Squared | 0.20 | 0.38 | 0.24 |

Note: OLS estimates with robust standard errors (in parentheses) are clustered at the subject level and OLS in column (1)-(2). Column (3)-(4) look at the difference in completion time and payoff when comparing the two treatments, controlling for the demographics variables including major, gender, age, whether received statistics training, experience with economics experiment, investment experiment, and use of GPT.. *** p<0.01, ** p<0.05, * p<0.1.

## Appendix F Experiment Robustness

### Appendix F.1 Motivation

The motivation behind this experiment is to address concerns that strategic uncertainty (Messick et al., 1988; Gangadharan and Nemes, 2009) may be misinterpreted as part of the cognitive uncertainty being measured.

The variable of interest, cognitive uncertainty, occurs when participants are unsure about their prosocial preferences or the utility-maximizing action in the public goods game that reflects their true social preferences. The main result in our study is that a higher cognitive uncertainty is coupled with a higher contribution. By contrast, strategic uncertainty arises due to unknown information about the decisions of others, i.e., individuals may be uncertain about the public good or resource requests in the common pool resource experiment as they do not know how other group members will behave[14]. One may argue that the uncertainty captured in this study is at least partly strategic uncertainty. Specifically, the existing literature has found that when manipulating strategic uncertainty by providing subjects with false information regarding the variance of contributions by others, a dramatic drop in contributions occurs when high environmental uncertainty[15] is combined with high strategic uncertainty (Wit and Wilke, 1998).

### Appendix F.2 Experimental Design and Procedure

There are 10 rounds of the game in the Experiment Robustness. At the beginning of the experiment, we randomly assign 1/4 of all subjects to the Full Information Treatment and remove strategic uncertainty for those subjects so that is not possible they misinterpret cognitive uncertainty. A screenshot for subjects who are in the Full Information are showed in Figure F.1.

Specifically, these subjects will always have access to the contributions of all other group members before making their own decisions, and they are informed that they will always be paired with three other subjects who are not in the Full Information Treatment. They are also informed in the instructions that the subjects who are not in the Full Information Treatment are unaware that they are playing with a subject who has access to their contribution decisions before they make

[14]Strategic uncertainty is sometimes known as social uncertainty. Literatures has pointed out that social uncertainty matters in the contribution, as the uncertainty regarding others' cooperation decisions would result in fear and greed and hence non-cooperative behaivor (Rapoport and Eshed-Levy, 1989) For example, when social uncertainty persists, subjects may fail to contribute to public goods because they fear that their contribution will be wasted, as the public goods will only be provided when the contribution exceed a certain threshold. Likewise, subject may also fail to contribute due to the greed to free ride others' contribution.

[15]Environmental uncertainty refers to the environmental variables that determine the optimal group action. In Wit and Wilke (1998) and Au (2004), they manipulate a high environmental uncertainty by raising the uncertainty about the provision threshold, i.e., the amount of contribution needed to provide the public good. They find that contributions decrease as environmental uncertainty increases. In contrast, there is no environmental uncertainty in all the experiment in our study.

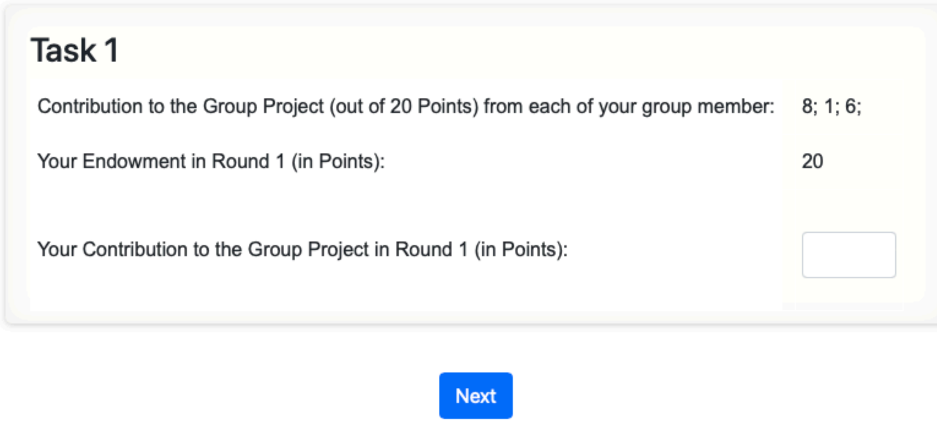



Figure F.1. Screenshot of an example decision screen in Treatment Full Information.

contribution decision.

We formulate the hypotheses for Experiment Robustness as follows:

**Hypothesis F.1:** *Cognitive uncertainty in Full Information is nonzero.*

**Hypothesis F.2:** *The correlation between cognitive uncertainty and contributions is statistically significant in the Full Information Treatment.*

### Appendix F.3 Experimental Result

A total of 84 subjects participated in the Experiment Robustness, with 21 subjects randomly assigned to the Full Information treatment.

The histogram of cognitive uncertainty in the two treatments is depicted in Figure F.2. When removing strategic uncertainty, there are approximately 12% more decisions report a zero cognitive uncertainty, and there is some evidence that the average cognitive uncertainty decreases when strategic uncertainty is removed ($t$ test: $t$=-0.0607, $p$=0.9517; Rank-sum test: $z = 2.031$, $p = 0.0423$). Nevertheless, the average cognitive uncertainty of 21.38, when strategic uncertainty is removed, is statistically significant from zero ($t = 10.1776$, $p = 0.0000$). This suggests that the cognitive uncertainty we measure in the main study is not entirely strategic uncertainty.

**Result F.1:** *Cognitive uncertainty is larger than zero even when strategic uncertainty is removed.*

Further, as shown in Table F.1, Columns (1)-(3), the observations from our main study are marginally supported with the 21 subjects sample whose strategic uncertainty is completely removed. First, there is a greater contribution from subjects with higher cognitive uncertainty (p=0.149), and their contributions are closer to half of the endowment when they exhibit more cognitive uncertainty. And once again, we fail to identify any inelastic pattern in the decisions made by subjects with cognitive uncertainty. Lastly, we do not find that subjects become more certain about their

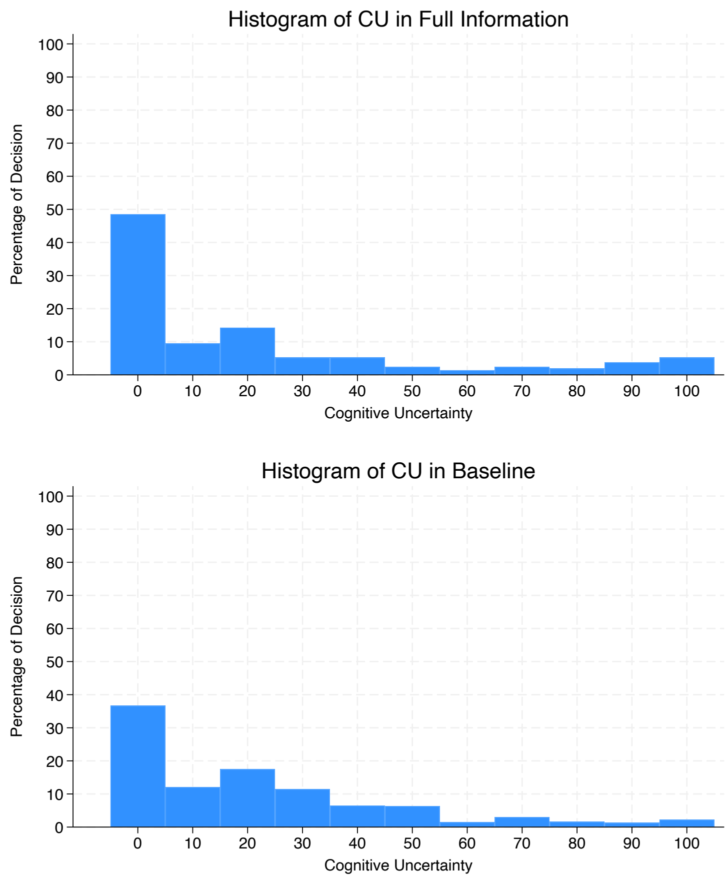



Figure F.2. Histogram of cognitive uncertainty

decisions as the game repeats, even in the absence of strategic uncertainty. Finally, subjects do not rely on past group members' contribution when they have access to this periods' data.

**Result F.2:** *We find evidence supporting the robustness of the results from our main study. When strategic uncertainty is eliminated, contributions are higher when subjects are more uncertain about their cognitive decisions.*

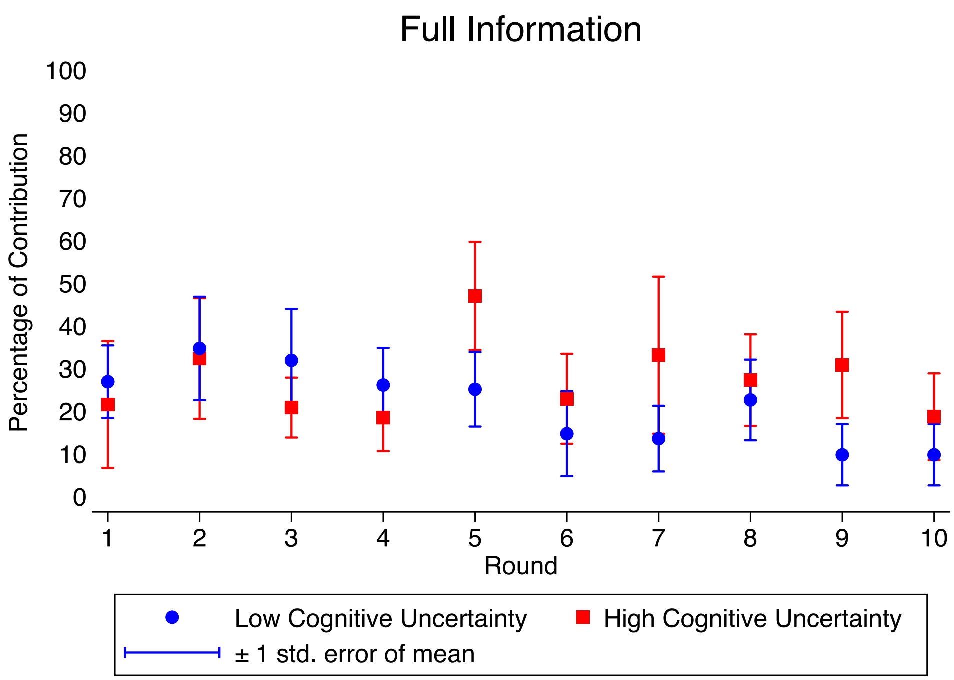



Figure F.3. Average contribution percentage to the public goods. Cognitive uncertainty is distinguished by comparing with the average of the cognitive uncertainty within a given round. High cognitive uncertainty decisions are decisions with cognitive uncertainty that are larger or equal to the average cognitive uncertainty within a given round, while low cognitive uncertainty is those with a cognitive uncertainty that are smaller than the average cognitive uncertainty within a given round.

Table F.1. Experimental Result in Full Information in Experiment Robustness

| Dependent Variable | Contribution<br><br>(1) | \|Contribution – 50% of Endowment\|<br>(2) | Contribution<br><br>(3) | Cognitive Uncertainty<br>(4) |
|---|---|---|---|---|
| Cognitive Uncertainty | 0.20 | -0.21*** | 0.12 | |
| | (0.13) | (0.06) | (0.19) | |
| Round | | | -2.00 | -0.05 |
| | | | (1.25) | (0.48) |
| Cognitive Uncertainty × Round | | | 0.01 | |
| | | | (0.02) | |
| others' % contribution | 0.23** | -0.11** | 0.17 | -0.04 |
| | (0.10) | (0.05) | (0.12) | (0.07) |
| Lagged members' % contribution | 0.15 | -0.03 | 0.10 | -0.06 |
| | (0.10) | (0.03) | (0.10) | (0.05) |
| Observations | 189 | 189 | 189 | 189 |
| R-squared | 0.08 | 0.12 | 0.66 | 0.71 |
| Number of Subjects | 21 | 21 | | |
| Approach | Subject FE Panel | | Subject FE OLS | |

Note. Column (1) -(2) report Fixed effects model with cluster-robust standard errors for panels nested within subject level. Column (3) – (4) reports the subject level fixed effect OLS estimates, with robust standard errors clustered at the subject level. *** p<0.01, ** p<0.05, * p<0.1