Four Essays in Repeated Games

Doctoral thesis for obtaining

the academic degree

Doctor of Economics (Dr. rer. pol.)

submitted by

Bao, Yongping

at the

Universität Konstanz

Faculty of Politics, Law and Economics

Department of Economics

Konstanz, 2023

Tag der mündlichen Prüfung:

18.07.2023

Namen der Referenten:

Sebastian Fehrler

Urs Fischbacher

Yves Breitmoser

Acknowledgements

I owe my heartfelt thanks to many people who have supported me throughout my PhD study. Without their encouragement and assistance, I could not have completed this journey within the past five years.

First and foremost, I would like to express my gratitude to my first supervisor, Sebastian Fehrler. I first met him during my Master's studies at Uni Konstanz, and he has been my guiding force ever since. Through working on several projects together, he has taught me the fundamentals of conducting research and has provided invaluable feedback not only on my papers and studies, but also on my future career plans. Additionally, he and his family have given me unwavering support during and after our move to Bremen, which I am truly grateful for. I would also like to extend my sincere thanks to my second supervisor, Urs Fischbacher. Under his mentorship, I have acquired numerous research techniques and gained valuable insights and inspiration from him and his chair at Uni Konstanz. I am also grateful to my third supervisor, Yves Breitmoser, for his valuable suggestions and feedback on my papers.

I would also like to express my deep appreciation for my friends and colleagues who have not only supported me, but also made my journey enjoyable during my studies at Uni Konstanz and Uni Bremen. I am especially grateful to Ankush Asri, Viola Asri, Regina Anselm, Deepti Bhatia, Fabian Dvorak, David Grammling, Angelica Serrano, Heiner Salomon for their unwavering support and companionship.

I would also like to extend my gratitude to my co-authors at Uni Konstanz and Uni Bremen who have provided me with invaluable support and insightful comments on our joint research projects. I am thankful to Ludwig Dandwitz, Hsuan Yu Lin, Volker Hahn, Lars Hornuf, and Bettina von Helversen for their valuable input.

During my time in Konstanz, I had the privilege of working with the Graduate School of Decision/Behavioral Science and the Thurgauer Wirtschaftsinstitut, both of which provided me with invaluable opportunities to conduct my research. I am deeply grateful for the chance they gave me to grow as a researcher. After moving to Bremen, I received immense support from the Center for Materials and Processes and from Socium, both of which generously provided funding to support my research and conference attendance. I am particularly grateful to Simone Ruiz Anderer for her invaluable assistance with administrative issues.

I would like to express my deepest gratitude to my family and close friends for their unwavering love and support throughout my academic journey. My parents' unconditional support and encouragement have been invaluable to me, and I am forever grateful for their belief in me. I also want to extend a special thanks to my best friend, Jie Luo, who has been a constant source of encouragement and support since high school. Without her, I would not have been able to achieve my academic goals.

Contents

Summary

This dissertation consists of four chapters that investigate human decisions in repeated games. All chapters present empirical results to the research questions using controlled laboratory experiments.

In the first chapter of this dissertation, we investigate algorithm-based decision-making in exploration and exploitation tasks. We utilize a reinforcement learning algorithm to advise human subjects in a multi-armed bandit task. To control for the algorithm's tendency towards exploration, we vary the weight it places on exploration, resulting in two algorithms: an explorative algorithm and an exploitative algorithm. We examine the number of times the subjects' choices align with the recommendations of the algorithm as a rough measure of their willingness to follow the algorithm's advice. In addition, we use cognitive modeling to estimate the subjects' latent tendency to follow the algorithm's recommendations. Our empirical results indicate that the algorithm's advice improves human decision-making. Furthermore, we find that the subjects' willingness to follow the algorithm heavily depends on the algorithm's latent exploration tendency. Specifically, subjects are more likely to follow an algorithm that is more exploitative, i.e., an algorithm that does not change its advised bandit frequency but is consistent in its recommendations, rather than following an algorithm that has a similar exploration tendency as the subjects themselves. Overall, our findings suggest that algorithm-based decision-making can be an effective tool to improve human decision-making, and that the design of the algorithm is crucial in determining its effectiveness.

The second chapter of this dissertation focuses on modeling an indefinitely repeated interaction between the public and a committee, such as the central bank. The research question is centered on public trust in the committee, which is essential for the functioning of many organizations. Theories suggest that committee structures can affect public trust in them. Specifically, we

distinguish between an individualist committee, which has a single decision-maker, and a collectivist committee, in which decisions are made collectively by a group of decision-makers. Within the collectivist structure, we further differentiate between a committee with synchronized terms and a committee with overlapping generational structure. Although theory predicts that trust is impossible to achieve under the individualistic structure, we find no difference in the trust rate empirically. This finding may arise from the mutual benefits of reciprocity. Therefore, trust in a committee is not solely dependent on the committee structure, but also on the underlying behavioral factors.

The third chapter of this dissertation focuses on cooperation in an indefinitely repeated Prisoners' Dilemma under imperfect public monitoring. In many economic interactions, players cannot directly observe the actual chosen actions of the other players, but instead receive noisy signals that represent those actions. Sustaining cooperation under these circumstances can be challenging. One important determinant of cooperation under imperfect monitoring is the correlation in the signals. When signals are correlated, players may exploit the correlation structure to infer the real actions of their opponents, making it easier to sustain cooperation. Our empirical findings, however, contradict the theoretical prediction that cooperation would be easier to sustain when signals are correlated. Specifically, our results show that subjects are equally cooperative whether the signals are correlated or independent. The reason for this unexpected finding is that when signals are correlated, subjects punish more frequently, as it becomes easier to detect defection. The reaction to defection becomes more severe, as there is no wiggle room surrounding bad signals.

The final chapter of this dissertation also focuses on cooperation in an indefinitely repeated Prisoners' Dilemma under imperfect public monitoring. Specifically, I investigate the impact of pre-play communication on cooperation under different levels of noise. When the noise is low, a cooperation equilibrium can be constructed, but as the noise level increases, cooperation becomes riskier. The central question I seek to answer is whether subjects will still cooperate even when it is theoretically impossible if they are allowed to communicate, and through what channel does communication affect cooperation. One possible mechanism through which communication could affect cooperation is by reducing the perceived strategic uncertainty. To measure perceived strategic uncertainty, I elicit subjects' beliefs about their opponents' actions. My findings show that the cooperation-sustaining effect of communication carries over to scenarios in which the noise level is high. Specifically, subjects who are allowed to pre-play communicate exhibit higher levels of belief in cooperation, which makes communication an effective means of mitigating the negative impact of noise on cooperation.

Zusammenfassung

Diese Dissertation besteht aus vier Kapiteln, die menschliche Entscheidungen in wiederholten Spielen untersuchen. Alle Kapitel präsentieren empirische Ergebnisse auf Forschungsfragen, die mit kontrollierten Labor-Experimenten untersucht werden.

Im ersten Kapitel untersuchen wir entscheidungsbezogene Algorithmen in Explorations- und Exploitationsaufgaben. Wir verwenden einen Reinforcement-Learning-Algorithmus, um menschlichen Teilnehmern in einer Multi-Armed-Bandit-Aufgabe Ratschläge zu geben. Um die Tendenz des Algorithmus zur Exploration zu kontrollieren, variieren wir das Gewicht, das er der Exploration beimisst, was zu zwei Algorithmen führt: einem explorativen und einem exploitativen Algorithmus. Wir untersuchen die Anzahl der Male, bei denen sich die Entscheidungen der Teilnehmer mit den Empfehlungen des Algorithmus decken, als grobes Maß für ihre Bereitschaft, dem Algorithmus zu folgen. Darüber hinaus verwenden wir kognitive Modellierung, um die latente Tendenz der Teilnehmer abzuschätzen, den Empfehlungen des Algorithmus zu folgen. Unsere empirischen Ergebnisse zeigen, dass die Empfehlungen des Algorithmus die menschliche Entscheidungsfindung verbessern. Außerdem stellen wir fest, dass die Bereitschaft der Teilnehmer, dem Algorithmus zu folgen, stark von der latente Explorationstendenz des Algorithmus abhängt. Insbesondere folgen die Teilnehmer einem Algorithmus, der stärker ausbeuterisch ist, also einen Algorithmus, der seine empfohlenen Banditenfrequenzen nicht ändert, sondern in seinen Empfehlungen konsistent ist, eher als einem Algorithmus, der eine ähnliche Explorationstendenz wie die Teilnehmer selbst hat. Insgesamt deuten unsere Ergebnisse darauf hin, dass entscheidungsbezogene Algorithmen ein effektives Instrument zur Verbesserung der menschlichen Entscheidungsfindung sein können und dass das Design des Algorithmus entscheidend für seine Effektivität ist.

Das zweite Kapitel dieser Dissertation konzentriert sich auf die Modellierung einer unbefristet wiederholten Interaktion zwischen der Öffentlichkeit und einem Ausschuss wie der Zentralbank. Die Forschungsfrage konzentriert sich auf das öffentliche Vertrauen in den Ausschuss, das für die Funktionsweise vieler Organisationen wesentlich ist. Theorien legen nahe, dass Ausschussstrukturen das öffentliche Vertrauen in sie beeinflussen können. Insbesondere unterscheiden wir zwischen einem individualistischen Ausschuss, der einen einzelnen Entscheidungsträger hat, und einem kollektivistischen Ausschuss, in dem Entscheidungen von einer Gruppe von Entscheidungsträgern kollektiv getroffen werden. Innerhalb der kollektivistischen Struktur unterscheiden wir weiter zwischen einem Ausschuss mit synchronisierten Amtszeiten und einem Ausschuss mit überlappenden Generationsstrukturen. Obwohl die Theorie besagt, dass Vertrauen unter der individualistischen Struktur unmöglich zu erreichen ist, finden wir empirisch keinen Unterschied im Vertrauensniveau. Dieses Ergebnis kann aus den gegenseitigen Vorteilen der Reziprozität resultieren. Vertrauen in einen Ausschuss hängt daher nicht nur von der Ausschussstruktur ab, sondern auch von den zugrunde liegenden Verhaltensfaktoren.

Das dritte Kapitel dieser Dissertation konzentriert sich auf die Kooperation in einem unendlich wiederholten Prisoner's Dilemma unter unvollständiger öffentlicher Überwachung. In vielen wirtschaftlichen Interaktionen können die Spieler nicht direkt die tatsächlich gewählten Aktionen der anderen Spieler beobachten, sondern erhalten stattdessen verrauschte Signale, die diese Aktionen darstellen. Die Aufrechterhaltung von Kooperation unter diesen Umständen kann eine Herausforderung darstellen. Ein wichtiger Faktor der Kooperation unter unvollständiger Überwachung ist die Korrelation in den Signalen. Wenn die Signale korreliert sind, können die Spieler die Korrelationsstruktur ausnutzen, um die tatsächlichen Aktionen ihrer Gegner zu erschließen, was es einfacher macht, die Kooperation aufrechtzuerhalten. Unsere empirischen Ergebnisse widersprechen jedoch der theoretischen Vorhersage, dass Kooperation einfacher aufrechtzuerhalten ist, wenn die Signale korreliert sind. Konkret zeigen unsere Ergebnisse, dass die Probanden unabhängig davon, ob die Signale korreliert oder unabhängig sind, gleich kooperativ sind. Der Grund für diese unerwartete Feststellung ist, dass die Probanden bei korrelierten Signalen häufiger bestrafen, da es einfacher ist, Abweichungen zu erkennen. Die Reaktion auf Abweichungen wird strenger, da es keinen Spielraum um schlechte Signale herum gibt.

Das letzte Kapitel dieser Dissertation beschäftigt sich ebenfalls mit der Kooperation in einem unendlich wiederholten Prisoner's Dilemma unter unvollständiger öffentlicher Überwachung.

Insbesondere untersuche ich die Auswirkungen der vorherigen Kommunikation auf die Kooperation unter verschiedenen Geräuschpegeln. Wenn das Rauschen niedrig ist, kann ein Kooperationsgleichgewicht konstruiert werden, aber je höher der Geräuschpegel ist, desto riskanter wird die Kooperation. Die zentrale Frage, die ich beantworten möchte, ist, ob Probanden trotz theoretischer Unmöglichkeit noch kooperieren werden, wenn ihnen die Kommunikation erlaubt ist, und durch welchen Kanal die Kommunikation die Kooperation beeinflusst. Ein möglicher Mechanismus, durch den die Kommunikation die Kooperation beeinflussen könnte, ist die Reduzierung der wahrgenommenen strategischen Unsicherheit. Um die wahrgenommene strategische Unsicherheit zu messen, erfrage ich die Überzeugungen der Probanden über die Aktionen ihrer Gegner. Meine Ergebnisse zeigen, dass der Kooperationserhaltungseffekt der Kommunikation auch auf Szenarien übertragen wird, in denen der Geräuschpegel hoch ist. Insbesondere zeigen Probanden, die vorherige Kommunikation haben, höhere Kooperationsüberzeugungen, was die Kommunikation zu einem effektiven Mittel macht, um die negative Auswirkung von Geräuschen auf die Kooperation zu mildern.

Chapter 1

Similarity and Consistency in Algorithm-Guided Exploration

With Fabian Dvorak, Ludwig Danwitz, Sebastian Fehrler, Lars Hornuf, Hsuan Yu Lin and Bettina von Helversen

Abstract

Algorithm-based decision support systems play an increasingly important role in decisions involving exploration tasks, such as product searches, portfolio choices, and human resource procurement. These tasks often involve a trade-off between exploration and exploitation, which can be highly dependent on individual preferences. In an online experiment, we study whether the willingness of participants to follow the advice of a reinforcement learning algorithm depends on the fit between their own exploration preferences and the algorithm's advice. We vary the weight that the algorithm places on exploration rather than exploitation, and model the participants' decision-making processes using a learning model comparable to the algorithm's. This allows us to measure the degree to which one's willingness to accept the algorithm's advice depends on the weight it places on exploration and on the similarity between the exploration tendencies of the algorithm and the participant. We find that the algorithm's advice affects and improves participants' choices in all treatments. However, the degree to which participants are

willing to follow the advice depends heavily on the algorithm's exploration tendency. Participants are more likely to follow an algorithm that is more exploitative than they are, possibly interpreting the algorithm's relative consistency over time as a signal of expertise. Similarity between human choices and the algorithm's recommendations does not increase humans' willingness to follow the recommendations. Hence, our results suggest that the consistency of an algorithm's recommendations over time is key to inducing people to follow algorithmic advice in exploration tasks.

1.1 Introduction

Algorithm-based decision support systems are becoming increasingly influential in the decision-making practices of individuals and organizations, including product searches, portfolio choices, and human resource procurement (e.g., Adomavicius and Tuzhilin, 2005b; Panniello et al., 2016; Tauchert and Mesbah, 2019; Zhou et al., 2021; Wang et al., 2022; Uotila et al., 2009). They also affect search and rescue missions and production settings, as well as geological and space exploration, in which teams of humans and algorithm-driven robots jointly perform search tasks (e.g., Orth et al., 2021; Schoonderwoerd et al., 2022; Pouya and Madni, 2021). Such decision problems are typically characterized by sequential choices and learning over time. The key trade-off is between searching longer for the best option, also referred to as *exploration*, or sticking to the option with the highest expected return at an earlier point in time, or *exploitation* (e.g., Addicott et al., 2017). People often struggle with such exploration–exploitation trade-offs due to the complexity of the tasks and constraints on cognitive resources (e.g., Gershman, 2020; Laureiro-Martinez et al., 2019). In contrast, it is widely recognized that reinforcement learning algorithms can perform very well in exploration–exploitation settings (Sutton and Barto, 2018), such as in the canonical multi-armed bandit problem (Thompson, 1933), making these tasks a compelling target for developing decision support systems (e.g., Zhou et al., 2021).

The usefulness and impact of advice from such algorithms crucially depends on the willingness of humans to follow it, and the literature on human–algorithm interactions provides little guidance in this regard. Previous research has primarily focused on one-shot decisions and the results have been fairly inconsistent. While in some judgment tasks, such as shopping online, humans happily follow an algorithm's advice (e.g., Zhou et al., 2021), in others they exhibit considerable reluctance to do so (Logg et al., 2019; Dietvorst et al., 2015). Indeed, recent reviews suggest that whether and when people are willing to take advice is complex, depending on the characteristics of the individual, the algorithm, and the task (Kawaguchi, 2021; Mahmud et al., 2022).

In this study, we investigate whether people accept and benefit from the advice of an algorithm in an exploration–exploitation task, and which characteristics of the algorithm and the human decision-maker influence the willingness to accept the algorithm's recommendation. In these tasks, the human decision-maker and the algorithm interact repeatedly, providing the opportunity to learn about the task but also allowing the human to observe the algorithm's decision behavior and quality of advice. The characteristics of the algorithm and the human decision-maker are therefore likely to jointly determine whether the advice is accepted and how effectively the task is solved. Given that the ability to balance exploration and exploitation is crucial to success in these tasks, we focus on whether the algorithm's tendency to explore and the similarity in tendencies between the algorithm and the human decision-maker will affect the latter's inclination to follow the algorithm's recommendations.

For this purpose, we conducted a tightly controlled online experiment using a stationary multi-armed bandit task—a typical task used to investigate exploration–exploitation trade-offs (Speekenbrink, 2022). Our analysis compares performance when human participants solve the task on their own versus when they receive advice from a state-of-the-art learning algorithm, which differs across treatments in its tendency to explore or exploit. We also use a cognitive modeling approach to estimate latent characteristics of the human decision-maker and compare these with the characteristics of the algorithm.

Our findings provide strong evidence that in a stationary multi-armed bandit task, participants benefit from the advice of a state-of-the-art learning algorithm independent of its exploration tendency. However, participants are more willing to follow an algorithm that is less explorative than they are. Participants that are more explorative than the algorithm benefit more from the algorithm's advice, which demonstrates the need to take the characteristics of both the human advisees and the algorithm into account when designing algorithm-based decision support systems.

The remainder of the paper is organized as follows: In Section 2, we provide a review of the literature, summarize the relevant theory, and explicate our research questions and hypotheses. Section 3 details our experimental method and design, and Section 4 presents the empirical results in line with our pre-registration and provides additional exploratory analyses. Section 5 discusses our findings and Section 6 concludes.

1.2 Related Literature and Behavioral Predictions

Experimental designs investigating how humans tackle exploration–exploitation trade-offs have frequently employed what are known as multi-armed bandit tasks (Daw et al., 2006; Gershman, 2020; Thompson, 1933; Sutton and Barto, 2018). In these tasks, the decision-maker must choose repeatedly between two or more options that differ in their expected rewards. Each choice results in probabilistic feedback drawn from an underlying reward distribution. In stationary multi-armed bandit tasks, the reward distribution across each of the options (or "bandits") is stable over time (for a recent review, see Speekenbrink, 2022). The decision-maker thus needs to balance exploration of previously unseen or rarely chosen options with exploiting options that have produced comparatively high rewards in the past. Except for a few special cases, it is impossible to compute optimal solutions to the trade-off between exploration and exploitation in the multi-armed bandit task (Schulz and Gershman, 2019). However, reinforcement algorithms have been shown to perform well in these tasks (Sutton and Barto, 2018) and usually outperform humans in experimental settings (Gershman et al., 2015).

One reason why people tend to underperform in exploration tasks is their reliance on random exploration rather than directed exploration. Random exploration refers to a noisy decision process, that is, one that involves choices that are not goal-directed and do not maximize rewards. Exploration is considered directed if it is undertaken to reduce uncertainty about the environment by exploring particularly informative options. Directed exploration thus leads to a preference for options that have only rarely been explored and thus are high in uncertainty (Wilson et al., 2021, 2014). Although humans have been shown to prefer exploring options with high uncertainty (Speekenbrink, 2022; Wilson et al., 2014; Wiehler et al., 2021; Zajkowski et al., 2017), their exploration behavior is often too random, particularly when cognitive resources such as working memory capacity are limited (Brown et al., 2022; Laureiro-Martinez et al., 2019; Meder et al., 2021; Wu et al., 2022). While algorithms are subject to similar limitations regarding working memory and speed, their constraints are often far less restrictive than those of humans. Hence, algorithm-based decision support systems with directed exploration have the potential to effectively help humans in exploration–exploitation tasks—if they are willing to take the advice of the algorithm.

Previous research on taking advice from algorithms has shown large differences in individuals' willingness to follow the recommendations of an algorithm (e.g., Kawaguchi, 2021; Logg et al., 2019; Mahmud et al., 2022). In general, humans seem to be willing to accept the advice of an algorithm if it is perceived to be of high quality and the expertise of the human is low (e.g., Logg et al., 2019; Saragih and Morrison, 2022; Tauchert and Mesbah, 2019; Van Swol et al., 2018). However, according to what Madhavan and Wiegmann (2007) refer to as the perfect automation scheme humans expect algorithms to work perfectly, unlike other humans, and adherence to an algorithm's recommendations decreases rapidly once the recommendations are perceived as imperfect. This can result in algorithm aversion in settings with stochastic environments, in which it is impossible to give perfect recommendations (Dietvorst et al., 2015; Dietvorst and Bharti, 2020; Prahl and Van Swol, 2017). Accordingly, one might expect algorithm aversion in a probabilistic task such as the multi-armed bandit task. In more recent research, however, participants have reported being likely to adopt an algorithm's advice when the algorithm's accuracy was above average or far greater than human performance (Saragih and Morrison, 2022). In addition, Filiz et al. (2021) have shown that algorithm aversion decreased when participants had the opportunity to evaluate the algorithm based on feedback over repeated decisions. Algorithms in multi-armed bandit tasks usually outperform human decision-makers (Gershman et al., 2015), particularly when participants have relatively little expertise in the task and receive feedback over the course of several trials. Therefore, we expect participants in general to follow the advice of a state-of-the-art learning algorithm. Accordingly, people should make more explorative choices when receiving advice from an explorative algorithm than when receiving advice from an exploitative algorithm. But given that both algorithms perform well in our parametrization, participants should benefit from the advice, independent of the exploration preferences of the algorithm.

Even though the algorithm's performance is important, it may not be the only factor influencing how much participants follow it. Most research on acceptance of the advice of algorithms has focused on one-shot decision problems, manipulating the expectations of advice quality by providing the advisee with information regarding the algorithm's past success (Hou and Jung, 2021). In contrast, exploration–exploitation tasks, such as a multiple-armed bandit task, involve a sequence of decisions in which the human decision-maker can repeatedly observe which options the algorithm suggests. The actual reward, however, is based on the participant's choice. On the one hand, this gives participants the chance to experience the quality of the algorithm's advice for themselves when following it. On the other hand, if participants are skeptical of the algorithm in the beginning and decide not to follow its advice, they may have fewer chances to obtain an accurate impression of the advice quality. Participants' initial perceptions of the algorithm, however, may depend less on the algorithm's ultimate success rate but rather on the characteristics of the algorithm, such as its exploration tendency. Thus our first research question is:

Research Question 1: Do people accept algorithmic advice in multi-armed bandit tasks? And if so, do they benefit from it?

Because multi-armed bandit tasks involve a trade-off between exploration and exploitation, algorithms can be designed to favor exploration or exploitation (Sutton and Barto, 2018) without necessarily affecting the algorithm's performance. This raises the question of whether people might be more likely to follow advice and reap greater benefit from an explorative or an exploitative algorithm.

Algorithms using directed exploration could be attractive as people have been found to show high levels of random exploration in multi-armed bandit tasks but struggle with directed exploration (i.e., strategic information-seeking) when cognitive resources are limited (e.g., Wu et al., 2022; Meder et al., 2021). Using directed exploration, explorative algorithms could help decision-makers to direct their natural tendency to explore options with high uncertainty and thus improve the information gained from exploration. Explorative algorithms are also likely to make suggestions that are less expected by the human decision-maker than exploitative algorithms. Research on recommender systems has identified novel and unexpected advice as a means to increase user satisfaction (e.g., Adamopoulos and Tuzhilin, 2014; Castells et al., 2022).

In contrast, algorithms with a stronger tendency to exploit are by design more consistent over time in their choices. In settings with sequential choices, consistency, as opposed to changing behavior or choices, could be perceived as a signal of expertise (e.g., Falk and Zimmermann, 2017; Fehrler and Hughes, 2018; Soll et al., 2022), which would make exploitative algorithms appear as having a higher ability. Indeed, Ihssen et al. (2016) found that in a reversal learning task, participants were more willing to follow choices that were consistent over time than choices that varied between options. In addition, the variance in the outcomes of choices recommended by exploitative algorithms is likely lower than the variance in outcomes of choices recommended by an explorative algorithm. Accordingly, choices recommended by exploitative algorithms will more likely result in rewards of average magnitude that are less likely to be perceived as a loss. They are also less risky. These aspects could increase the participants' willingness to follow the advice of an exploitative algorithm.

In sum, there are arguments for both preferences—for following a more explorative or a more exploitative algorithm—and our study helps to clarify the relative value of these styles.

Research Question 2: Does the willingness to follow the algorithm's advice depend on the algorithm's exploration tendency?

Lastly, the exploration tendency of the algorithm on its own may matter less than how well it corresponds to the human it provides advice to, as is reflected in the rising importance of personalization in recommender systems (e.g., Adomavicius and Tuzhilin, 2005a; Adomavicius et al., 2008). Research on human decision-making in exploration–exploitation tasks suggests that people exhibit stable tendencies toward either exploration or exploitation in multi-armed bandit tasks (von Helversen et al., 2018; Zettler et al., 2020). Furthermore, research on advice-taking indicates that people are more willing to follow advice of people that they perceive as similar to them, for instance because they share the same gender, come from the same geographical region (Gino et al., 2009), or share similar personality traits (Tauni et al., 2019). In addition, similarity between the advice and one's own judgment in the absence of advice can impact the willingness to accept advice (Minson et al., 2011; Schultze et al., 2015). Advice that differs strongly from the advisee's own independent judgment is often ignored, even when following the advice would enhance the recipient's performance (Ecken and Pibernik, 2016; Yaniv, 2004). Similarly, Capponi et al. (2022) find that clients are especially willing to follow investment advice from a robo-advisor if it adapts to the investor's risk profile, which also results in better investment strategies. These results suggest that people may be more willing to take advice from an algorithm with an explorative tendency similar to their own. Accordingly, we hypothesize that the difference between the exploration tendency of the person and the exploration tendency of the algorithm will influence the person's willingness to follow the advice.

Research Question 3: Does the similarity between the exploration tendency of the algorithm and the human decision-maker influence the latter's willingness to accept the advice?

1.3 Experimental Design and Method

1.3.1 Study Overview

To investigate our research questions, we implemented a ten-armed stationary bandit task in an online experiment, in which the participants' goal is to maximize their rewards. Participants' payments depend on the number of points they have earned during the task, either on their own or with the help of advice from a Kalman filter algorithm with exploration bonus (KAE). The Kalman filter algorithm is a state-of-the-art reinforcement learning algorithm (Chakroun et al., 2020; Daw et al., 2006) adjusted to solve stationary bandit problems, and is particularly suited to this research given its strong performance and its ability to model human learning and choice behavior in multi-armed bandit tasks. In particular, the algorithm has been shown to better capture human behavior than comparable models, such as delta learning models (Chakroun et al., 2020; Speekenbrink and Konstantinidis, 2015). We use the same Kalman filter algorithm to provide advice to participants in the experimental setting and to *ex post* model the human decision-process, which enables us to measure similarity in the exploration behavior based on the algorithm's exploration parameter. For this purpose, the individual explorative tendency of the participants estimated by the model from the experimental data is compared with the explorative tendency of the algorithm.

The experiment consists of two phases. In the first phase, all participants perform the multi-armed bandit task without advice, which allows us to estimate the participants' own tendency to explore. In the second phase, we implement four treatment conditions between participants: three treatments in which participants perform the multi-armed bandit task with advice from an algorithm, and a control treatment in which participants perform the task without advice, to ensure that potential increases in performance are due to the advice and not due to learning. In the treatments with advice, the algorithm gives recommendations on which alternative to choose, but leaves it to the human participants to make the final decision. The exploration tendency of the algorithm varies by three levels: explorative, exploitative, and balanced. The algorithm in the explorative treatment has a stronger tendency toward exploration. In the exploitative treatment, the algorithm is more inclined toward exploitation. In the balanced treatment, the algorithm's tendency to explore lies between the explorative and exploitative treatment. Note, that the algorithm provides its outline of recommendations based on the participant's actual choices and the algorithm's tendency to explore. Thus, our experimental design captures a true interaction between a human and an algorithm.[1]

1.3.2 Participants

Participants were recruited through Prolific Academic. Participation was restricted to UK residents to ensure good knowledge of English and comparable value of the study incentives. Participants had to correctly answer comprehension questions about the task goal, reward scheme, and reward generation procedure in order to participate in the experiment. Participants who failed the comprehension questions were removed from the experiment, and replaced by new participants until the pre-registered number of participants was reached. Six hundred participants finished the study. Among these, thirteen have been be excluded from the data analyses: twelve because their data was not uploaded due to technical issues, and one participant because he indicated that he had failed to read the instruction regarding the algorithm. Participants' payments consisted of a base payment of £2.50 and a performance-dependent bonus payment. The average bonus was £4.14 per participant. Of the participants willing to provide their demographic details, 65.1% are female, 34.2% are male, and 0.7% are non-binary. The average age is 40.8, with a standard deviation of 23.4 years.

1.3.3 Bandit Task and Experimental Procedure

In the experiment, participants perform a total of eight instances of a stationary multi-armed bandit task with 10 bandits: one training task and three experimental tasks each in the first and the second phase of the experiment. In each task, participants encounter ten boxes (the bandits) and must undertake several trials. In each trial, participants may choose one box by clicking on it with a mouse. After selecting a box, its contents are revealed and the participant receives feedback about the number of points collected from the box. The goal is to collect as many points as possible, and participants have been told that their payout depends on the

[1] The main hypotheses and analyses were preregistered on the Open Science Foundation (OSF) website, `https://osf.io/e3hqa`.

average number of points they accumulate over the course of the experiment. The boxes differ in their underlying reward distributions, whose means have been randomly sampled from a normal distribution with a mean of 50 points and a standard deviation of 10. Each time a box is selected, a random number is drawn from a normal distribution with a mean of the average points of the box and again a standard deviation of 10. The drawn number is then rounded to the closest integer. For a visualization of an example reward distribution, see Figure 1.1, Panel A. A training task consists of 20 trials and a experimental task consists of 70 trials. Within a task, the average points that a box will yield stays the same, that is, within an experimental task participants have 70 choices to learn which boxes result, on average, in high payoffs and which result in low payoffs. For each new task, the average points for each box are redrawn. During a task, participants are informed about the current number of trials and the maximum number of trials in the task. After each task, they are informed about the average number of points they have acquired during the task.

Participants begin the experiment with a declaration of consent informing them about the processing of their data and the university conducting the experiment (University of Bremen). After providing consent, participants receive detailed instructions explaining the task, the experimental procedure, and the reward scheme.[2] In the first phase of the experiment, participants perform one training task followed by three experimental tasks without advice from an algorithm. In the second phase of the experiment, participants again perform one training task followed by three experimental tasks. Depending on the treatment participants are in, they either perform the training and experimental tasks on their own (control) or, at the beginning of the phase, read additional instructions regarding the algorithm's advice, before receiving these recommendations throughout the training and experimental tasks. The advice depends on the treatment and is provided by either an explorative, balanced or exploitative algorithm. The recommendation of the algorithm is indicated by a yellow square surrounding the suggested box. Participants are informed that the algorithm has received the same information as the participants, that is, the algorithm is aware of the distribution of the average bandit points and the standard deviation of each draw from a bandit. The algorithm is also informed of the number of points from the bandits picked by the participants. The recommendations from the algorithm are made based on this information.

[2]See Appendix A.8 for the consent form and the instructions.

After the participants finish the second phase, they are informed about the average points they have acquired over the course of the experiment and the amount of the bonus payment they will receive. Participants are then asked several questions regarding their subjective perceptions of the experiment. Specifically, they are asked whether there are any reasons to disregard their responses, and to rate the effort they put into the experiments and their tendency to rely on exploitation and exploration throughout the experiment. If the participants belong to the treatments with algorithm recommendations, they are also asked about the usefulness of the recommendations, how much they relied on the algorithm, and how much they followed the algorithm, each using a slider that records their responses between 0 (not at all) and 1 (very much). All the participants are given the opportunity to leave comments regarding the experiment before ending the experiment.[3]

1.3.4 Algorithm Recommendations

To generate the recommendations, we use a Kalman filter algorithm with exploration bonus (KAE) (Chakroun et al., 2020; Daw et al., 2006). In each trial $t \in \{1, \ldots, 70\}$, the algorithm uses Bayesian learning to generate posterior expectations about the mean and standard error of the mean of the points generated by each bandit. Let $\hat{\mu}_{k,t}$ denote the posterior expectation of the mean, and $\hat{\sigma}_{k,t}$ the posterior expectation of the standard error of the mean of bandit k in trial t. In the first trial, the expected means and standard errors correspond to the prior expectation, i.e., $\hat{\mu}_{k,1} = 50$ and $\hat{\sigma}_{k,1} = 10$ for all bandits. In all subsequent trials, the algorithm updates the expected mean and standard error of the mean of the bandit selected in the current trial based on the Kalman filter (Daw et al., 2006). The Kalman filter uses the following three recursive equations to update the posterior expectations of the selected bandit s conditional on the reward R_t obtained by this bandit in trial t:

$$\begin{aligned} \hat{\mu}_{s,t} &= \hat{\mu}_{s,t-1} + K_t(R_t - \hat{\mu}_{s,t-1}) \\ \hat{\sigma}^2_{s,t} &= (1 - K_t)\hat{\sigma}^2_{s,t-1} \\ K_t &= \frac{\hat{\sigma}^2_{s,t-1}}{\hat{\sigma}^2_{s,t-1} + \sigma^2_{s,0}} \end{aligned} \tag{1.1}$$

The quantity K_t is called the Kalman learning rate, which is updated based on the expected variance of the selected option. As the expectation for the standard error of the mean equals the

[3]The experiment was programmed in PsychoPy (Peirce et al., 2019) and exported to Pavlovia.

known standard deviation of the average rewards before the bandit is sampled, and subsequently decreases, the learning rate starts at one-half and decreases over trials. This implies that the expectation of the standard error of the mean decreases every time the bandit is sampled, which is a specific feature of stationary bandits.

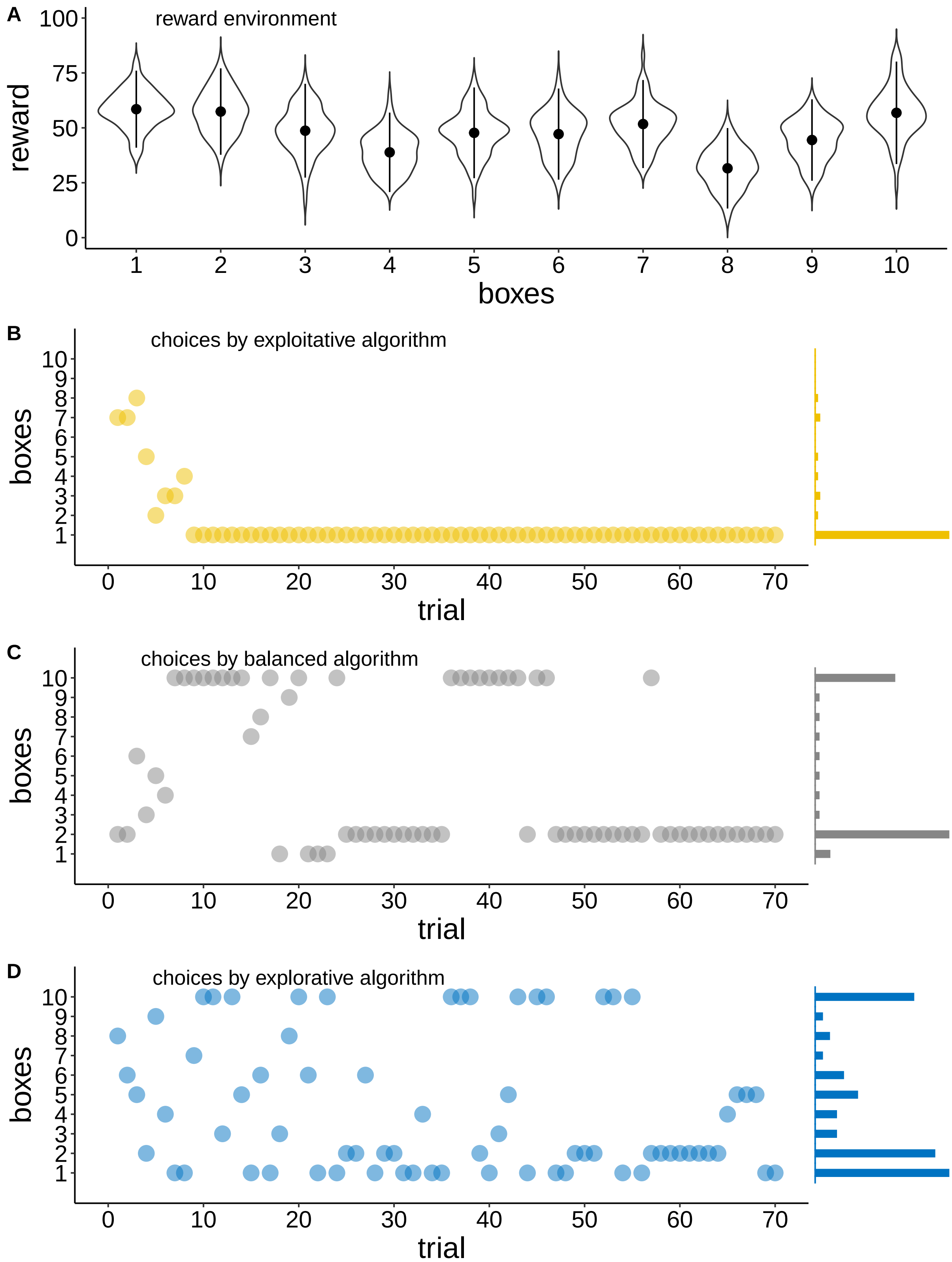
The algorithm recommendation in trial t depends on the posterior expectations $\hat{\mu}_{k,t}$ and $\hat{\sigma}_{k,t}$, and a parameter ϕ that reflects the algorithm's tendency toward exploration. The recommendation for trial t is the bandit with the largest value of the attraction score $q_{k,t} = \mu_{k,t} + \phi\sigma_{k,t}$. If there are multiple bandits sharing the largest attraction score, the bandit with the larger s will be recommended.

The parameter ϕ controls the exploration tendency of the algorithm by scaling the effect of the estimated standard error of the mean on a bandit's attraction score. If ϕ is large, bandits with a large standard error of the mean (and which are less frequently chosen) have large attraction scores. If ϕ is small, uncertainty does not attract, and the algorithm will mostly recommend bandits based on the posterior expectations about their mean reward.

We use three different values of the exploration parameter ϕ across the three treatments with algorithm recommendations. In the exploitative treatment $\phi = 0.4$. The balanced treatment uses $\phi = 1.4$, and the explorative treatment $\phi = 3.3$.

The three values of ϕ have been selected based on their performance in 10,000 simulated tasks of our experiment. The values used in the exploitative and explorative treatments both generate the same average reward of approximately 61 points across the 10,000 simulated tasks. These values have been selected because they generate an average reward comparable to Bayesian learning in combination with Thompson sampling, which is known to perform well in multi-armed bandit tasks (Sutton and Barto, 2018). The value of $\phi = 1.4$ that we use in the balanced treatment produces an average reward of 62 points across the 10,000 simulation runs, which suggests an optimal balance between exploration and exploitation. Panels B–D of Figure 1.1 illustrate how the advice by the algorithms in the advice treatments differs by presenting an exemplary sequence of choices when the algorithms perform the task without the human participant.

Figure 1.1: Exemplary Algorithm Choices



Note: Panel A depicts an exemplary reward environment. When facing this environment alone, i.e., with no human participant involved, the exploitative algorithm produced the choice sequence depicted in Panel B, the balanced algorithm produced the choice sequence depicted in Panel C, and the explorative algorithm produced the choice sequence depicted in Panel D.

1.4 Results

First, we check whether the different algorithms' exploration tendencies in our experimental treatments influence participants' behavior. Following the preregistration, we test whether the algorithm's recommendations influence participants' exploration behavior, the rewards they earned, and how frequently their choices match the recommendation.[4] Second, we report the cognitive modeling results and how participants' exploration tendencies affected their willingness to follow the algorithm's advice, as well as participants' subjective perceptions of the algorithm's usefulness.

1.4.1 Behavioral Analyses

Exploration Tendency: Number of Switches.

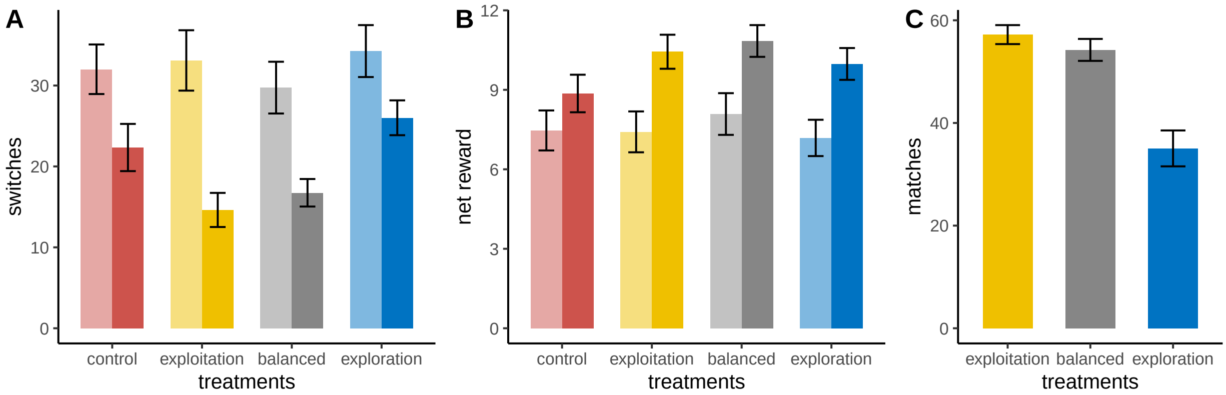
Participants' exploration tendency is quantified by measuring how frequently they switched between bandits within each of the two experimental phases, averaged across tasks, with a higher number of switches indicating a higher exploration tendency.

By summarizing the average number of switches per task in the two phases, Panel A of Figure 1.2 shows that participants in general explored more frequently in the first phase ($M = 32.28$, $SD = 20.47$) than in the second phase ($M = 19.94$, $SD = 14.62$; $F(2, 1172) = 141.3$, $p < .001$, partial $\eta^2 = .11$; see also Table 1.1).

To test whether the characteristics of the algorithm influenced participants' exploration behavior, we analyze whether treatment differences affect the number of switches in phase 1 and 2 differently. The ANOVA for phase 1, with no recommendations in all treatments, shows no differences between the four treatments ($F(3, 583) = 1.31$, $p = .271$, partial $\eta^2 = .007$), showing that initial exploration tendencies of the participants were balanced across treatments. In phase 2, both an ANOVA including all four treatments and one focusing only on the three recommendation treatments show a significant effect of treatment ($F(3, 583) = 20.47$, $p < .001$, partial $\eta^2 = .10$ and $F(2, 439) = 35.42$, $p < .001$, partial $\eta^2 = .14$, respectively), indicating that the recommendations changed how much participants explored.

[4]The Figures in Appendix A.6 report changes in behavior over time, i.e., across tasks.

Figure 1.2: Summary of Behavioral Measures by Treatment and Phase.



Note: Panel A shows the average number of switches, Panel B the average net rewards minus the expected value of 50, and Panel C the average percentage of matches between participants' choices and the algorithm's suggestions. All measures are averaged across the three tasks in phase 1 (light colors) and phase 2 (dark colors), respectively. Error bars indicate 95% confidence intervals.

Table 1.1: Descriptive Overview of Behavioral Measures by Treatment

	treatments	phase	M	SD	95% lower bound	95% upper bound
number of switches	control	1	32.01	18.82	28.94	35.07
	exploitation	1	33.10	23.00	29.37	36.83
	balanced	1	29.74	19.85	26.54	32.94
	exploration	1	34.26	19.92	31.05	37.47
	control	2	22.34	17.92	19.42	25.25
	exploitation	2	14.62	12.99	12.51	16.73
	balanced	2	16.75	10.55	15.05	18.45
	exploration	2	26.01	13.35	23.86	28.17
rewards	control	1	57.47	4.64	56.71	58.22
	exploitation	1	57.41	4.75	56.64	58.18
	balanced	1	58.09	4.89	57.30	58.87
	exploration	1	57.18	4.26	56.50	57.87
	control	2	58.86	4.35	58.15	59.57
	exploitation	2	60.44	3.96	59.80	61.08
	balanced	2	60.84	3.72	60.24	61.44
	exploration	2	59.98	3.73	59.38	60.58
number of matches	control	-	-	-	-	-
	exploitation	2	57.21	11.40	55.37	59.06
	balanced	2	54.23	13.27	52.09	56.37
	exploration	2	35.04	21.70	31.55	38.54

To investigate the differences in exploration behavior between treatments more closely, we conducted post-hoc pairwise comparisons between the treatments using Tukey HSD corrections for multiple tests.[5] The pairwise comparisons show that participants in the exploitation treatment (M = 14.6, SD = 13.0) explore less than participants in the exploration treatment (M = 26.0, SD = 13.4; $exploit - explore = -11.39$, $p < .001$), but not compared to participants in the balanced treatment (M = 16.7, SD = 10.6; $balance - exploit = 2.12$, $p = .559$). Participants in the exploration treatment explored more than in the balanced treatment ($balance - explore = -9.27$, $p < .001$). In comparison to the control treatment (M = 22.3, SD = 17.9), participants explored less in both the exploitation and the balanced treatment ($exploit - control = -7.72$, $p < .001$; $balance - control = -5.60$, $p = .004$), but there was no difference in the exploration condition, ($explore - control = 3.68$, $p = .110$).[6]

Result 1: Participants' exploration behavior depends on the exploration tendency of the algorithm. When participants are presented with an exploitative algorithm, they tend to explore less (exploit more) than when they are presented with an explorative algorithm or when exploring on their own.

Rewards.

Next, we turn to the question of whether participants' performance improves when receiving recommendations from the algorithm. Panel B of Figure 1.2 shows the net rewards (i.e., the average number of points per task minus the expected value of 50) participants achieved during the two phases. Overall, participants' performance increased from phase 1 (M = 57.54. SD = 4.64) to phase 2 (M = 60.03, SD = 4.00; $F(1.1172) = 97.37$, $p < .001$, partial $\eta^2 = .08$, see also Table 1.1).

In the first phase, in which participants have not yet received advice from an algorithm, we find no significant differences between the treatment conditions, ($F(3, 583) = 1.02$, $p = .382$, partial $\eta^2 = .005$), as expected. In phase 2, however, net rewards differ significantly depending on the four treatment conditions, ($F(3, 583) = 6.87$, $p < .001$, partial $\eta^2 = .03$). Importantly, this difference seems to stem from lower performance in the control condition, given that an

[5]See Tables A.5.1-A.5.3 in Appendix A.5 for more details. In Appendix A.7, we show the empirical cumulative distributions of switches in the two phases across treatments.

[6]Using Bonferroni adjustments instead of the Tukey corrections leads to comparable results (details not reported here).

ANOVA focusing on the three recommendation treatments without control does not show a significant difference between these treatments ($F(2, 439) = 1.93$, $p = .147$, partial $\eta^2 = .009$).

Again we conducted follow-up paired comparisons between the treatments using the Tukey HSD correction. In comparison to the control treatment ($M = 58.9$, $SD = 4.35$), subjects earned higher rewards in the exploitation treatment ($M = 60.4$, $SD = 3.96$; $exploit - control = 1.58$, $p = .004$) and the balanced treatment ($M = 60.8$, $SD = 3.72$; $balance - control = 1.98$, $p < .001$), but not in the exploration treatment ($M = 60.0$, $SD = 3.73$; $explore - control = 1.12$, $p = .074$). However, the three treatments with the AI did not differ significantly from each other ($exploit{-}explore = 0.46$, $p = .749$; $balance{-}explore = 0.87$, $p = .233$; $balance{-}exploit = 0.41$, $p = .814$).[7]

Result 2: Participants earn greater rewards when receiving advice from an exploitative or balanced algorithm than when performing the task on their own.

Matches.

Next, we look at how frequently participants' choices coincided with the recommendations from the algorithm, as a rough measure of how willing participants are to follow the algorithm's recommendations. As Panel C of Figure 1.2 illustrates, the three treatments differ in the average frequency with which participants chose the bandit that was recommended by the algorithm ($F(2, 439) = 82.26$, $p < .001$, partial η^2=.27, see also Table 1.1).

Specifically, follow-up Tukey tests with HSD corrections show that participants' choices match the algorithm's recommendations less frequently in the exploration treatment than in the other two treatments ($exploit{-}explore = 22.17$, $p < .001$; $balance{-}explore = 19.19$, $p < .001$). However, comparing the balanced treatment with the exploitation treatment reveals no significant difference in the frequency of matches between these two treatments ($balance{-}exploit = -2.98$, $p = .252$). Overall, the results suggest that participants seem to be more willing to follow the algorithm's recommendation when the algorithm is exploitative or balanced than when it is explorative.

[7]Independent sample t-tests with Bonferroni corrections for multiple hypothesis testing confirm the result, showing a significant advantage for the exploitation and the balanced treatment, but not for the exploration treatment over the control treatment (adjusted p-values are $p = .004$, $p < .001$ and $p = .095$, respectively).

Result 3: Participants' decisions coincide less frequently with the algorithm's recommendations when the algorithm has a more explorative tendency.

To investigate whether the frequency of matches depends on tendency to explore, we regressed the number of times a participant followed the algorithm's recommendations on the difference between the participant's exploration behavior in phase 1 (the number of switches in tasks 1–3) and the algorithm's recommended exploration behavior in phase 2 (the average number of recommended switches by the algorithm in tasks 4–6) and the squared difference using OLS regression controlling for age, gender and treatment conditions.[8] The regression shows that the recommendations of the algorithm are followed more frequently if the algorithm is less explorative than the participant (see Tables A.2.1 and 1.2 in Appendix A.2). We also find a smaller but significant effect of the squared difference, which suggests a curve-linear relationship. Moreover, when splitting the data set into participants that are more or less explorative than the algorithm, as shown in Figure 1.3, we find that the tendency to make the same choice declines when the algorithm is more explorative than the participant but not when the algorithm is more exploitative than the participant (see Table 1.2 for the corresponding regression results).

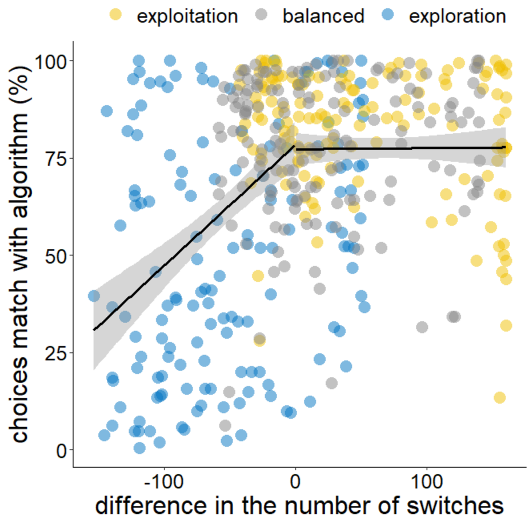
Result 4: Participant behavior is more likely to coincide with an algorithm's decisions when the algorithm is more exploitative and therefore more consistent over time than the participant.

1.4.2 Cognitive Modeling

A potential issue with behavioral measures of exploration and algorithm following is that these measures do not control for the variation in rewards observed by the participant. For example, in situations in which one bandit stands out due to large rewards, even a participant with a strong preference for exploration will rarely switch. For the same reason, a participant might often follow the recommendations of the algorithm in the second phase of the experiment simply because recommendations and preferences are very similar, which makes it difficult to judge

[8]We also performed a one-way ANOVA with 1,000 iterations as a balance check on age and gender. The p-values are .464 for age and .495 for gender, indicating again that the participants are well randomized into the treatments.

Figure 1.3: Scatter Plot of the Difference in the Tendency toward Exploration of Participants and the Algorithm with Matching



Note: The figure shows two regression lines of percentage of matching on difference in the number of treatments. We look at difference below and above zero. The regression lines use all observations from non-control treatments. The difference in the number of switches is the participant's number of switches in phase 1 minus the algorithm's treatment-wise average recommended number of switches in phase 2.

the latent tendency to follow recommendations based on the frequency of following. Thus, to corroborate the behavioral results, we use cognitive modeling to estimate each participant's latent tendency toward exploration as well as the latent tendency to follow the recommendations of the algorithm.

To estimate the latent exploration tendency of each participant, we fitted a model with a random exploration parameter β to the first-phase choices. The model uses the same Bayesian learning algorithm we used to generate the recommendations, which returns an attraction score $q_{s,t}$ for each bandit based on the posterior expectations about the mean and standard deviation of the points generated by the bandit. Instead of using a deterministic choice rule, we use the *softmax* function with inverse temperature β to translate the attraction scores of the bandits into stochastic choices to account for decision errors of the participants. The probability of participant i to select bandit s in trial t is:

$$Pr(s) = \frac{e^{\beta_i q_{s,t}}}{\sum_k e^{\beta_i q_{k,t}}} \tag{1.2}$$

Table 1.2: Effects of Dissimilarity and Algorithm Consistency on Matching

	without consistency			with consistency		
	Est	SE	p-value	Est	SE	p-value
constant	164.14	5.14	$< .001$	216.07	3.98	$< .001$
I(more explorative)	-11.36	5.86	.053	-2.12	3.84	.582
dissimilarity	-0.86	0.05	$< .001$	-0.17	0.04	$< .001$
I(more explorative) × dissimilarity	1.01	0.07	$< .001$	0.18	0.06	.001
consistency	-	-	-	-0.75	0.03	$< .001$
N	442			442		
R^2	0.55			0.81		

Note: Linear regressions of number of matches on a dummy variable (*I(more explorative)*) that indicates whether the participant explores more than the algorithm; the *dissimilarity* in the exploration tendency, which is is the absolute difference between the number of switches of participants in phase 1 and the number of recommended switches from the algorithm in phase 2; and the interaction of the dummy and dissimilarity. The second model additionally controls for the *consistency* of the algorithm, i.e., the number of switches recommended by the algorithm in phase 2. Age and gender are used as socio-demographic control variables.

We use Bayesian multilevel modeling to estimate the inverse temperature parameter β_{i1} for each participant in the first phase of the experiment.[9] To assure that the inverse temperature is positive, we use $\beta_1 = e^x$ and a normally distributed prior for the underlying variable x. We report the natural logarithm $log(\beta_1)$ of the inverse temperature in the first phase of the experiment to obtain an approximately normally distributed variation measure of participants' exploration tendencies.

We also fitted the same model to the choices participants make in the second phase of the experiment (tasks 4–6), which yields a parameter estimate β_2 for each participant. For the data of the control condition, comparing β_1 to β_2 allows us to assess the stability of the latent exploration tendency over the course of the experiment. For the treatments with algorithm recommendations, the model fits provide benchmarks for answering the question of whether accounting for the algorithm recommendations increases out-of-sample prediction accuracy.

To make reasonable comparisons between the exploration tendency of a participant and the exploration tendency of the algorithm, we also fitted the model to the algorithm recommendations. We did this on the individual level for each participant separately to obtain the algorithm's exploration tendency β_{alg} from the perspective of each participant. We use $log(\beta_1)–log(\beta_{alg})$ as

[9]For brevity, we drop the subscript i from here onward. Keep in mind however, that β_1 takes different values between participants.

a measure of difference between the exploration tendency of a participant and the algorithm, with positive values indicating that the participant has a stronger tendency toward exploitation.

We also fitted a two-parameter model with a random exploration parameter and a directed exploration parameter to the data of the first part of the experiment. The individual parameter estimates show that the variation in exploration behavior as measured by the number of switches stemmed from different degrees of random exploration. Participants generally avoided directed exploration, with little individual variation in this parameter. This can be attributed to our use of stationary bandits that discourage exploration in the later trials of each task.

For the treatment conditions with recommendations, we use a different model for the second-phase choices. In this model, we do not estimate the random exploration parameter for each individual. Instead, we fixed β to the mean of the posterior parameter distribution of each participant that we estimated based on their first-phase choices. This allows disentangling of the effect of algorithm recommendations from participants' exploration tendencies. To this end, we added a dummy $I_{s,t}$ to the equation of the attraction scores indicating whether the bandit was recommended by the algorithm, which yields:

$$q_{k,t} = \mu_{s,t} + \rho_i I_{s,t-1} \tag{1.3}$$

The parameter ρ_i of this dummy reflects the participant i's inclination to follow the recommendations of the algorithm. Positive values of ρ increase the probability that the recommended bandit is chosen. As in the model of first-phase choices, attraction scores are translated into choice probabilities using a *softmax* with inverse temperature fixed to the estimate of β_1.

The models were fit to the choice data using the Bayesian model fitting software Stan (Carpenter et al., 2017), accessed via rstan (Stan Development Team, 2022). For each treatment condition (exploration, balanced, exploitation, control) we fitted two Bayesian multilevel models: one for the first-phase choices and another for the second-phase choices. To approximate the posterior distribution of model parameters, we run four Markov Chains parallel with 3,500 iterations, of which 1,000 iterations were discarded as warm-up. Using this procedure, we obtain converged chains (max(Rhat) = 1.049) with informative posterior distributions (min(ESS) = 162) and no divergent transitions. Past research indicates that the models used meet the standards of parameter and model recovery (Danwitz et al., 2022).

For the population average of β, we use a log-normal prior with mean 0.5 and standard deviation of 0.1, and normal distribution with mean 0 and standard deviation of 0.01 censored at zero as prior for standard deviation of β in the population. For the model parameter ρ, we use a standard normal prior for the population mean and a censored standard normal distribution for the standard deviation of ρ in the population. Using these priors, we conducted a parameter recovery simulation showing that individual model parameters are recovered well from simulated data (see Figure A.3.1 in Appendix A.3). Table A.4.1 in Appendix A.4 reports the posterior of the population mean and standard deviation of the model parameter for each fitted model, along with criteria for convergence and estimates of the out-of-sample prediction accuracy of each model (Vehtari et al., 2017). Comparing the out-of-sample prediction accuracy of the ρ models to the β models for the data of second-phase choices indicates that modeling choices based on recommendations improves out-of-sample prediction accuracy (exploitation: $z = -12.05, p < .001$; balanced: $z = -33.19, p < .001$; exploration: $z = -47.57, p < .001$).

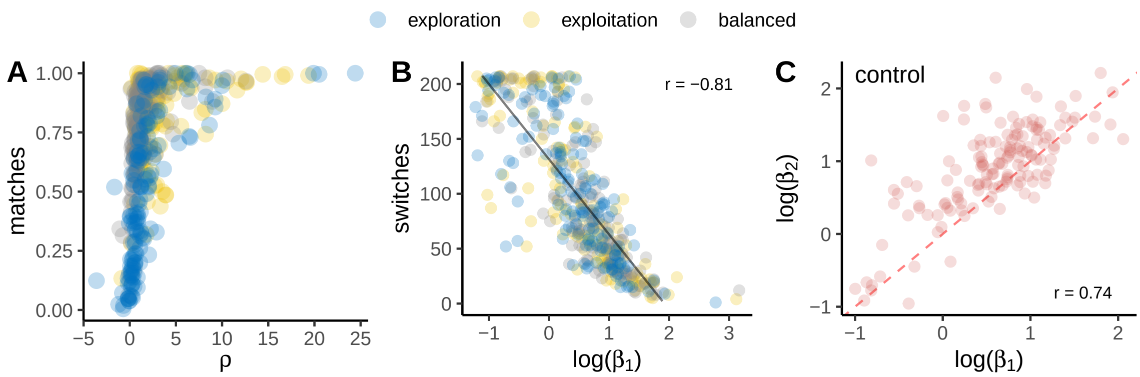
Analyses of Participants' Latent Tendencies toward Exploration and Inclination to Follow the Algorithm.

As a first step, we examine whether the estimated latent variables capture individual differences in the behavioral measures. As illustrated in Figure 1.4, the parameters show high correlations with the corresponding behavioral measures. Panel A shows the relationship between the parameter ρ, which measures the latent tendency to follow recommendations, and the number of times a participant chooses the bandit recommended by the algorithm in the second part of the experiment. Both variables are positively correlated, with Pearson's $r = 0.43$. The scatter plot also clearly shows the *softmax* function the model uses to translate recommendations into choice probabilities, with values of $\rho > 3$ resulting in more than 90% matches on the behavioral level. Panel B shows that the latent exploration tendency, as measured by $log(\beta_1)$, is negatively correlated with the number of switches in the first phase of the experiment. The larger a participant's latent exploration parameter, which indicates more exploitation and less exploration by the algorithm, the less frequently a participant switches between bandits (Pearson's $r = -0.81$, $t(440) = -28.87$, $p < .001$). Panel C of Figure 1.5 illustrates that the latent exploration tendency is fairly stable over the two phases of the experiment for data of the control treatment (Pearson's $r = 0.74$, $t(143) = 13.00$, $p < .001$). There is a slight tendency toward less exploration in the second phase, indicated by the fact that most observations are above the 45° line.

Result 5a: The latent measures of a participant's exploration tendency and willingness to follow an algorithm correlate highly with the corresponding behavioral measures.

Result 5b: Participants' latent exploration tendencies are relatively stable over time.

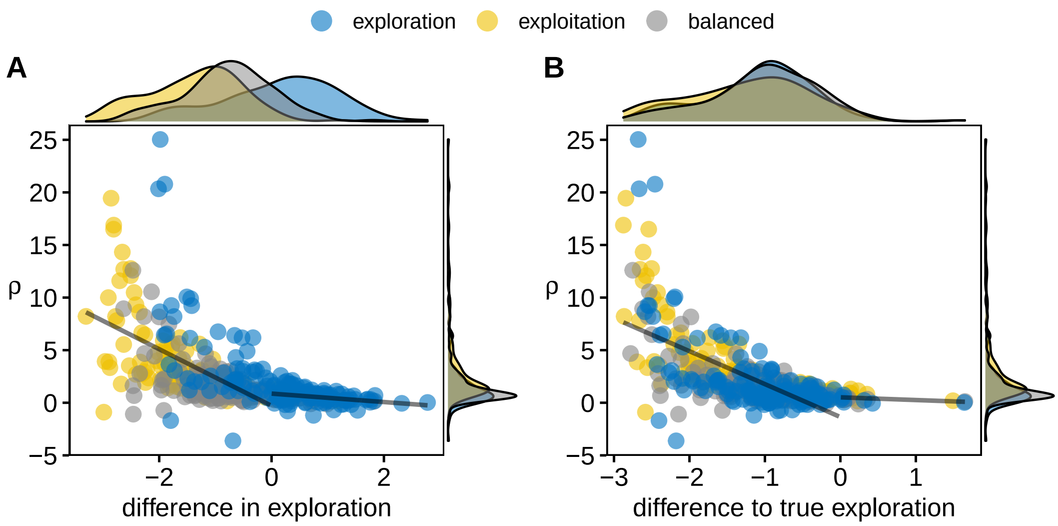
Figure 1.4: Correlations between Parameter Estimates and Behavioral Measures



Note: Panel A shows the relationship between the latent tendency to follow recommendations ρ and the average frequency of choosing the recommended bandit in the second phase. Panel B shows the latent exploration tendency (higher $log(\beta_1)$ corresponds to less exploration) and the number of switches in the first phase. Panel C shows the correlation between the estimated latent exploration tendencies of the first and second phase of the control treatment.

Corresponding to the behavioral analyses, we use the latent parameters, estimated based on the cognitive modeling, to investigate how the similarity between the participants' and the algorithms' exploration tendencies relates to participants' inclination to follow the algorithm's advice. As illustrated in Panel A of Figure 1.5, we find a negative relationship between the estimated latent tendency to follow the recommendations of the algorithm and the difference in the exploration tendency of the participant and the algorithm. We use $log(\beta_1)$–$log(\beta_{alg})$ as a measure of difference between the exploration tendency of a participant and the algorithm, with negative values indicating that the participant explored more in phase 1 than the algorithm recommended in phase 2. The latent tendency to follow recommendations is largest for participants who explored more than the algorithm recommends. The tendency to follow the recommendation is substantially lower if the exploration tendencies of the participant and the algorithm were similar and stay at a very low level when the exploration tendency of the algorithm was higher than the participants' exploration tendency. On average, the tendency to follow recommendations is largest in the treatment with the exploitative algorithm, which also features the most cases in which the algorithm was more exploitative than the participants.

Figure 1.5: Scatter Plot of the Differences in Exploration Tendencies and the Inclination to Follow the Algorithm



Note: The figure shows the relation of the parameter estimates of the latent tendency to follow recommendations ρ and the difference in the exploration tendency of the participant and the algorithm, with higher values indicating stronger exploitation (i.e., less exploration) by the participant than the algorithm. Panel A compares the participants' exploration tendency with the perceived exploration tendency of the algorithm ($log(\beta_1)$–$log(\beta_{alg})$) and Panel B with the uninfluenced exploration tendency of the algorithm ($log(\beta_1)$–$log(\beta_{exo})$. Colors indicate the different treatment conditions with recommendations. Parameter distributions are shown on the opposite axes respectively.

Given that the algorithm uses the feedback from the choices of a participant to generate a recommendation, participants' choice behavior can influence the degree to which the algorithm recommends exploration or exploitation and thus the exploration tendency displayed by the algorithm. We estimate the uninfluenced exploration tendency of the algorithm based on simulated recommendations and compare it with the exploration tendency estimated from the recommendations in the second phase. To this end, we simulated the recommendations the algorithm would have given if the participant had followed all recommendations and fitted the model for the first-phase choices to the simulated recommendations. The resulting exogenous exploration tendency β_{exo} allows us to study how the difference between the participant's and the algorithm's true exploration tendency $log(\beta_1)$–$log(\beta_{exo})$ is related to the latent tendency to follow the recommendations. Panel B of of Figure 1.5 shows a very similar negative, non-linear relationship between the estimated latent tendency to follow the recommendations of the algorithm and the difference between the participant's and the algorithm's unbiased exploration tendency. Table A.4.2 in Appendix A.4 shows that the non-linear relationship between the tendency to follow recommendations and the difference between the participant's and the algorithm's unbiased exploration tendency prevails in all treatments when controlling for socio-demographic variables.

We also test whether the relationships (depicted in the two panels of Figure 1.5) are significant when controlling for consistency of recommendations, gender, and age. To this end, we regress the individual estimates of ρ on a dummy that indicates whether the participant explores more than the algorithm $I(log(\beta_1) < log(\beta_{exo}))$, the dissimilarity in the exploration tendency $|log(\beta_1)$–$log(\beta_{exo})|$, and their interaction, controlling for the consistency $log(\beta_{exo})$ of the algorithm. Table 1.3 shows that the dissimilarity between the exploration tendency of the participant and the algorithm is positively related to the inclination to follow the algorithm recommendations if the exploration tendency of the participant is stronger. This corresponds to the linear fits depicted for the data points below zero in Figure 1.5. For this data, the dissimilarity has a significant positive effect on ρ in both models, even when controlling for the consistency of the recommendations, which shows that relative differences in the exploration tendency matter.

Result 6: Participants are more inclined to follow the algorithm if it is less explorative and therefore more consistent over time than they are.

Biased Perceptions and Benefits of Recommendations.

In order to characterize how participants' choices bias the level of exploration exhibited by the algorithm, we conducted exploratory analyses that go beyond our preregistration and focus on the absolute bias $log(\beta_{alg}) - log(\beta_{exo})$ introduced by the choices of the participant. Panel A of Figure 1.6 relates the absolute bias induced by a participant's choices to the difference in the exploration tendency of the participant and the algorithm. Biases are mainly negative, which means that the participant's perception of the algorithm is distorted toward more exploration. The scatter plot in Panel A of Figure 1.6 shows that differences in exploration explain the bias

Table 1.3: Effects of Dissimilarity and Algorithm Consistency on ρ

	diff exploration						diff true exploration					
	Est	SE	p-value	Est	SE	p-value	Est	SE	p-value	Est	SE	p-value
constant	0.72	0.42	0.084	1.92	0.43	$<.001$	0.25	0.62	0.690	1.96	1.16	0.091
I(more explorative)	-1.25	0.45	0.006	-0.38	0.45	0.391	-1.95	0.63	0.002	-1.92	0.63	0.003
dissimilarity	-0.41	0.41	0.319	-1.56	0.42	$<.001$	-0.39	0.99	0.693	-0.47	0.99	0.637
I(more explorative) ×dissimilarity	3.21	0.46	$<.001$	4.75	0.48	$<.001$	3.62	1.01	$<.001$	3.74	1.01	$<.001$
consistency	-	-	-	-1.75	0.24	$<.001$	-	-	-	-1.12	0.64	0.081
N	442			442			442			442		
R^2	0.40			0.46			0.46			0.46		

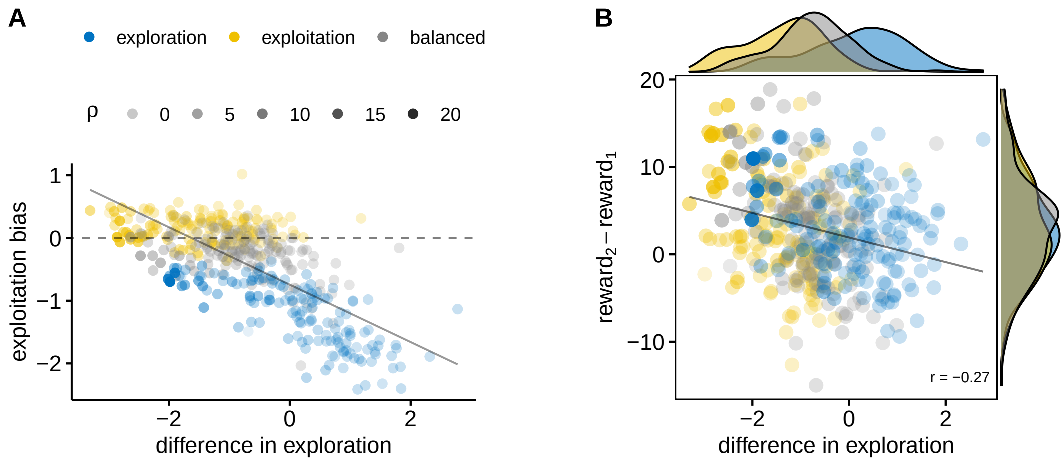
Note: Linear regressions of the latent tendency to follow algorithm recommendations, on a dummy that indicates whether the participant explores more than the algorithm (*I(more explorative)*), the *dissimilarity* in the exploration tendency $|log(\beta_1)–log(\beta_{alg})|$ of the participant and the algorithm, the interaction of the dummy, and dissimilarity. The second models additionally control for the consistency $log(\beta_{alg})$ of the algorithm. The right columns show the same models using $log(\beta_{exo})$ instead of $log(\beta_{alg})$. Age and gender are used as socio-demographic control variables.

(Pearson's $r = -0.72$, $t(440) = -21.66$, $p < .001$). It further shows that positive differences in exploration, which are mainly observed in the exploration treatment, correspond to negative biases in the perception of the algorithm, which gives the impression that recommendations are more random. In contrast, the perception bias is smaller for participants with negative differences in exploration, who are also those that follow recommendations more frequently (see Figure 1.5).

Result 7: The perception of the algorithm's behavior is biased in the direction of more exploration by interacting with a human participant. This bias is particularly strong when the algorithm is more explorative than the human.

Based on the individual parameter estimates, we can also answer the question of who benefits most from recommendations. Panel B of Figure 1.6 plots the differences in average reward over the two phases of the experiment against the difference in the exploration tendency of the participant and the algorithm. Most participants gain more in the second phase of the experiment with the help of recommendations, independent of the treatment condition. Panel B also illustrates that participants who explore the most also benefit the most from receiving recommendations (Pearson's $r = -0.27$, $t(440) = -5.86$, $p < .001$). As pointed out above, these participants are also the ones who exhibit the largest latent tendency ρ to follow the recommendations.

Figure 1.6: Biased Algorithm Perceptions and Heterogeneous Benefits



Note: A: Difference in the exploration tendency $log(\beta_1)–log(\beta_{alg})$ of the participant and the algorithm (positive values indicate more exploration by the algorithm than the participant) and bias in the perception of the exploration tendency of the algorithm $log(\beta_{alg})-log(\beta_{exo})$ (negative values indicate a bias toward more perceived exploration). B: Difference in exploration tendency and difference in reward between the first and the second phase of the experiment.

Result 8: Participants that have a high tendency to explore are most likely to follow the advice of an algorithm and also benefit most from receiving the advice.

1.4.3 Humans' Subjective Perceptions of the Algorithm

At the end of the experiment, we asked participants how useful they found the algorithm's advice, how much they relied on the algorithm and how much they followed the algorithm's advice. The answers to these three questions are sufficiently highly correlated (all $rs > .72$) that we have combined them into a single index of the perceived helpfulness of the algorithm. The subjective ratings are in line with the cognitive modeling results. The rating of the algorithm is positively correlated with participants' latent inclination to follow the algorithm ($r = .59$, $p < .001$). Participants also rated the helpfulness of the algorithm higher in the exploitative (M=.67, $SD = 0.21$) and balanced treatment (M=.65, $SD = 0.25$) than in the explorative treatment (M=.45, $SD = 0.29$, see Appendix A for more details).

Result 9: Subjective perceptions of the algorithm's usefulness correspond to the estimated inclination to accept its advice. They are in line with the findings from the cognitive modeling analysis as well as with the behavioral results.

1.5 Summary of Results and Discussion in Light of the Literature

1.5.1 Summary of Results

This study examines how human behavior in a stationary bandit task changes in response to receiving advice from algorithms with different exploration tendencies. In line with our expectations, we find that participants explore more when they obtain recommendations from a highly explorative algorithm than when the recommendations come from a balanced or more exploitative algorithm. In both the low-exploration and the balanced treatment conditions, participants explored less than in the control treatment, while participants in the control and in the high-exploration treatment conditions were on the same exploration level. Moreover, the participants exhibited the ability to use the algorithm's recommendation to enhance their performance: in all treatments with algorithmic advice, participants earned higher net rewards than in the control condition, while there was no significant difference among the treatments with recommendations. These results show that even in a probabilistic task, such as our multi-armed bandit, people are willing to accept advice from an algorithm and by doing so improve their performance. These results are in line with research showing that when an algorithm performs well (Saragih and Morrison, 2022) and people have the opportunity to gain experience with it, they are willing to follow it (Filiz et al., 2021).

We also sought to test whether participants' willingness to follow the algorithm's advice depended on the exploration tendency of the algorithm and the similarity between it and their own exploration tendency. We hypothesized that the similarity of the advising algorithm and the participants regarding the inclination to explore impacts the usage of advice. This was operationalized in two ways: analyses of behavioral data and a cognitive modeling approach.

The two analyses reveal a similar pattern. On the behavioral level, we found that for those participants who individually would explore more than the algorithm would recommend, matches between participants' choices and the algorithm's suggestions increase with similarity. For those participants who would individually explore less than the algorithm recommends, there is no distinct relationship between switches and matches. Similarly, the cognitive modeling analyses show that participants' latent inclination to follow the algorithms' advice does not increase with similarity, but that participants are more willing to follow an algorithm's advice the more exploitative the algorithm is when compared to their own exploration tendency. While participants with a high exploration tendency often follow recommendations that are less explorative than they are, participants follow the advice when the algorithm suggests more exploration less frequently. Moreover, participants' subjective ratings of the algorithm's usefulness align with the behavioral and modeling analyses. Thus, presumably, following or not following algorithmic advice is a rather deliberate process.

In sum, we find no evidence that a similarity in exploration tendencies of the algorithm and the human advisee increases the willingness to follow the advice, but rather that the willingness to follow depends on the exploration preferences of the algorithm. Even though the quality of the advice, i.e., the algorithm's success rate, was equal, participants were only willing to frequently accept the advice when the algorithm was less explorative and thus gave more consistent advice over time. One reason could be that consistency is regarded as a signal of expertise (Falk and Zimmermann, 2017; Fehrler and Hughes, 2018), which is supported by participants' reports of a higher usefulness of the algorithm's advice in the exploitative and balanced treatments. Following this line of reasoning, participants may have hesitated to follow recommendations to explore because they interpreted exploration as randomness or a demonstrated lack of expertise. This line of reasoning confirms previous findings that humans lose their trust in algorithms rapidly once they observe algorithms committing errors (Dietvorst et al., 2015).

As pointed out above, most participants nevertheless enhanced their performance when they received algorithm recommendations. On average, participants in the explorative treatment also benefited from the advice. Our analyses show that this may be the case because even in the explorative condition some participants exhibited a stronger exploration tendency than the algorithm and were willing to follow it. Indeed, participants who individually explore extensively not only followed the recommendations the most, but also benefited the most from the recommendations in terms of their earned net rewards. This shows that not only the exploration tendency of the algorithm but also the exploration tendency of the participant should be considered, and that highly explorative participants are the most likely to benefit from an algorithm's advice in tasks with exploration–exploitation trade-offs.

Lastly, our explorative analyses show that interacting with a human decision-maker can affect the behavior of the algorithm as the algorithm uses the participants' outcomes to inform its recommendations. The less explorative participants are in our setting, the stronger the algorithm's perceived bias toward giving explorative recommendations. While this effect is found in all three algorithm implementations, it is most prevalent for the high-exploration algorithm and least prevalent for the low-exploration algorithm. Continuously ignoring the algorithm's recommendation to explore provoked the algorithm to advise more exploration, which likely decreased participants' willingness to follow the algorithm even further. This should be taken into account in the design of algorithms for practical use cases.

1.5.2 Discussion in Light of the Literature

Our results inform two major debates in the current literature on taking advice from algorithms: (1) the debate on whether humans underutilize, rely too much on, or make appropriate use of algorithm recommendations, and (2) the debate on the importance of adviser–advisee similarity. On the practical side, our results also inform the development of algorithms to provide advice in decision support or recommender systems.

Although our approach does not include a control condition with human instead of algorithmic recommendations, it still informs the debate on algorithm aversion (Dietvorst et al., 2015; Dietvorst and Bharti, 2020; Logg et al., 2019). In their review on algorithm aversion, Jussupow et al. (2020) mainly identify the performance level of algorithms and information regarding their performance level as determinants of algorithm aversion or algorithm appreciation. In our setting, participants are biased toward the low-exploration algorithms, in relation to their individual tendencies, despite the fact that the low-exploration and high-exploration algorithms are equally good and participants obtain the same information regarding both. Hence, further investigations on algorithm aversion should take into account whether the algorithm under investigation is explorative or exploitative. If the asymmetry of participants' responses to low-exploration recommendations generalizes to different tasks and scenarios, it might systematically affect the exploration of teams of humans and algorithms: if humans rely on low-exploration recommendations while ignoring high-exploration recommendations, this will bias the tendency of human–algorithm teams away from exploration. This needs to be taken into account by the designers of such algorithms.

On the debate over the importance of similarity between adviser and the advisee, which has been mainly studied in settings where humans take advice from other humans (Yaniv, 2004), the question whether such findings generalize to taking advice from algorithms is gaining interest (Schemmer et al., 2022; Pálfi et al., 2022; Himmelstein, 2022). Previous studies suggest that humans are more likely to accept advice from other humans that are similar to them or give similar advice (e.g., Gino et al., 2009; Minson et al., 2011; Tauni et al., 2019; Yaniv and Kleinberger, 2000). In the present study, we do not find any evidence that similarity in decision strategies increases willingness to accept an algorithm's advice. Future research is necessary to disentangle whether these findings stem from the nature of the advice giver, from the type of similarity considered, or from the sequential nature of our task as compared to one-shot judgment tasks. One indication of the importance of the task is Ihssen et al.'s (2016) finding that participants were more likely to follow a human with a more consistent strategy in a reversal learning task. This suggests that consistency is a key factor in participants' willingness to follow in tasks with sequential choices more generally.

Decision support systems often deal with well known, structured environments, for example, when clinicians diagnose diseases (Ganju et al., 2020). When applying decision support systems to situations where less information is available and trial-and-error learning might occur, such as in geological exploration or supply chain management, decision support systems are faced with the exploration–exploitation trade-off (Chaharsooghi et al., 2008). It has been shown that decision support systems can be especially beneficial for human performance in such less-predictable environments (Van Bruggen et al., 1998). However, our research indicates that it is not only important how the decision support system addresses this trade-off, but also how its chosen strategies account for the human factor. The concept of giving explorative recommendations has structural similarities to the concept of unexpected recommendations (Adamopoulos and Tuzhilin, 2014). Such recommendations that take the receivers' expectations into account have been shown to enhance the performance of recommendation systems. However, in our setup, we find the opposite effect.

Zhou et al. (2021) evidence that e-commerce product-search algorithms providing more refined, i.e., exploitative, recommendations increase the usefulness of the search algorithm. This is in line with our finding that recipients of algorithm recommendations value recommendations that are exploitative and consistent over time. Zhou et al. (2021) report that this effect is more pronounced for those people who already know what product they want. Exploratory consumers, on the other hand, engage more with the platform when search algorithms provide them with a broader range of recommendations. Similarly, in our research we can pinpoint the individual's exploration tendency as a reference point for the participants' perception of the algorithm. However, it is likely that our highly structured task gives rise to less curiosity and serendipity (Shani and Gunawardana, 2011) and therefore participants' exploration tendency and their interest in explorative recommendations might be lower than they would be in richer and less structured exploration settings.

1.6 Conclusion

We set out to study the effects of algorithmic advice, and the willingness of advisees to follow such advice, in a tightly controlled experimental exploration task. Our results show that participants benefit from the recommendations of algorithms in structured exploration–exploitation environments. When the algorithms are more exploitative than the advisee, the inclination to follow the algorithm increases, as does the advisee's subjective perception of the helpfulness of the algorithm. And while it is important to take the exploration tendency of the advisee into account, our research does not support the notion of a *per se* effect of similarity. Even though algorithmic recommendations cannot, and do not, always lead to good outcomes in every trial of the exploration of a probabilistic environment, most people are willing to accept the advice, unless it looks too explorative to them.

These results contributes novel insights to the rapidly expanding literature on how best to design algorithmic decision support systems (e.g., Adomavicius and Tuzhilin, 2005b; Panniello et al., 2016; Tauchert and Mesbah, 2019; Wang et al., 2022; Uotila et al., 2009). The importance of the consistency of advice relative to the advisee's own inclination toward exploration, which our results highlight, could be further investigated in future research on real customer recommendation systems, such as the one explored by Zhou et al. (2021).

Chapter 2

Public Trust in Organizations

With Sebastian Fehrler and Volker Hahn

Abstract

Public trust is crucial for the functioning of many organizations, such as central banks. Therefore, it is important to understand the institutional characteristics affecting them. More specifically, we ask whether an individualistic or a collectivist structure makes an organization more trustworthy and whether communication in the form of organizational mission statements increases public trust. To address these questions, we study repeated versions of a basic trust game in which the trustee is an organization where decisions are either made by an individual or a collective. A game-theoretic analysis implies that public trust may or may not occur for a collectivist structure with overlapping terms but that public trust is impossible to achieve for an organization dominated by an individual. Empirically, we find that, even when the person in charge changes frequently, an individualistic organization can achieve similar levels of cooperation as a long-lived collectivist organization. This may indicate that cooperation is mainly driven by intrinsic motivation rather than by a desire to influence the future decisions of others.

2.1 Introduction

The success of most organizations depends on the trust that outsiders have in them. Central banks can achieve low and stable inflation much more effectively if citizens are confident that the central bank will be successfully stabilizing inflation in the future.[1] The performance of other agencies such as the police, public administration, or commercial banks also relies to a large extent on the trust that citizens or customers have in them. For example, depositors' concerns about the stability of a financial institution can lead to bank runs, which puts the survival of these institutions at risk. Therefore, it is crucial to examine which institutional structures help inducing public trust.

Specifically, we focus on two institutional dimensions. The first dimension concerns whether the organizational structure is individualistic or collectivist. Important decisions in an organization may be made by either an individual or a group of decision-makers. While most central banks have committees that take monetary policy decisions, they are located somewhere between an individualistic and a collectivist structure, as, in most cases, the president assumes a particularly powerful and prominent role.[2] In an individualistic organization, turnover may make it difficult to adopt policies that are desirable for both parties in the long term but costly in the short term. This may result in low levels of trust. On the other hand, turnover may also create opportunities to re-establish trust when past levels of trust were low. Compared to an individualistic organization, a collectivist organization provides more continuity, which may foster cooperation and trust.[3] However, the diffusion of responsibilities among multiple decision-makers may also make the organization opaque and hence less trust-worthy.

Second, we focus on the potential role of announcement of an official strategy. Many organizations have mission statements. Relatedly, central banks typically announce official strategies such as a specific variant of inflation targeting. These statements may play a dual role. They serve as a guideline for decision-makers within the organization and influence their behaviors. These statements also signal the objectives of the organization to outsiders, which may be conducive to trust.

[1] This is a central implication of the new Keynesian model, where current inflation depends on expectations about future inflation. See Woodford (2003) for a textbook treatment.

[2] For example, it is common to label different periods of U.S. monetary policy after chairpersons of the FOMC ("Volcker Disinflation", "Volcker-Greenspan era", "pre-Volcker era",...).

[3] Kroszner and Stratmann (2000), for example, argue that committees facilitate interactions and thereby allow for reputation development in repeated interactions between legislators and interest groups.

In this study, we propose a simple framework to examine how these two dimensions affect public trust in organizations. In particular, we study repeated versions of the canonical trust game (Berg et al., 1995) with binary choices. The public (the trustor) can send a share of its endowment to an organization (the trustee). If it does, the organization receives an additional surplus that captures the gains from cooperation. The decision-makers in the organization can then decide to send a given fraction of the received transfers back to the first player. In this case, the public receives these transfers and an additional surplus. Thus, the public's beliefs about whether the organization will reciprocate the transfers are crucial for the public's decision to submit transfers to the organization in the first place. Trust facilitates cooperation in our framework, where by cooperation, we mean that both the public and the organization transfer to one another, which maximizes aggregate payoffs.

In the variant of our model that focuses on an individualistic organization, this (stage) game is embedded into a framework where a player representing the public interacts indefinitely often with an organization in which all decisions are made by a single individual with a fixed term in office. After the end of a decision-maker's term, she is replaced by a new decision-maker with the same fixed term length. The model variant formalizing the collectivist organization is identical, except that all decisions are made by a committee of overlapping generations of decision-makers with fixed individual terms. The player representing the public can observe the decisions of the organization but not the behaviors of the individual committee members.

We then examine the implications of this model, both theoretically and experimentally. In the case of an individualistic organization, standard backward-induction arguments imply that the public never considers the organization trustworthy in the sense that it never transfers funds to the organization. This lack of trust is justified, as the organization would never reciprocate transfers. In the case of a collectivist organization, no unique equilibrium is obtained, even if we impose the refinement that non-pivotal players vote sincerely. Interestingly, there are equilibria with full cooperation; that is, the public always transfers its endowment to the organization, which then reciprocates the transfers. These equilibria implement Pareto-efficient allocations. However, in a collectivist organization, there are also equilibria in which no cooperation occurs and no transfers are made. Our theoretical predictions are thus not clear-cut.

This is one of the motivations for conducting an experiment. The second motivation is that cooperation occurs in experiments and even in one-shot interactions. The trustor transfers funds to the trustee based on the belief that they will reciprocate (see, for example, Berg et al., 1995).[4] Hence, one might expect positive transfers and cooperation even in an individualistic organization. Moreover, from a theoretical perspective, cooperation is feasible only if the terms of members overlap but not if all members' terms end simultaneously.[5] However, in a lab experiment, Xu and Potters (2018) find overlapping terms are only mildly conducive to cooperation between members of an organization. Hence, it is instructive to examine whether an overlapping term structure has an effect under a collectivist organization. Therefore, our experimental analysis distinguishes between two different treatments of collectivist organizations: one with overlapping terms and one with synchronized terms.

We find that the average rate of trust is high in all treatments. The average trust and cooperation rates do not differ between treatments. Individualistic organizations allow for the same levels of trust as collectivist organizations when when the interactions are relatively short-lived. This suggests that repeated interactions among the same individuals are of minor importance for cooperation behavior. We further find a last-term effect in individualistic organization but not in collectivist organizations in which the terms are overlapping. The rate of cooperation is high until the last term before the decision-maker leaves the committee. Many decision-makers in the organizations seem to reciprocate the public's trust for future benefits until the last term. The overlapping structure helps stabilize trust and cooperation over time. However, in the individualistic treatment, despite a decrease in the rate of sending in the last term, the average sending rate of the committee member is still quite high. This may imply that a high-share of decision-makers are reciprocates who reciprocate even without futrue gains. Therefore, trust in a committee is not solely dependent on the committee structure, but also on the underlying behavioral factors.

As previously mentioned, we consider a second factor that potentially influences trust in an organization: mission statements. Examining mission statements is particularly relevant for a collectivist organization with overlapping terms, as mission statements may help players coordinate on Pareto-superior equilibria. The initial members of the organization are allowed to select one of the two pre-specified statements about the choices that the organization should

[4] In finitely repeated gift-exchange game, subjects cooperate for both reciprocal and reputation concerns (Gächter and Falk, 2002). In our design, as the members of an organization interact with the public for fixed term-length, we might as well observe a repeated game effect that cooperation is established until the decision-maker's last term in the office.

[5] To be more precise, this statement is true only if we impose a refinement akin to trembling-hand perfection, which rules out equilibria where decision-makers vote against their interests as they are never pivotal.

make. One statement says that the organization will always cooperate, while the other statements says the opposite. If these announcements affect cooperation rates, it will be interesting to examine whether this effect disappears after some time when all initial members have left, or whether the statements can also have a more long-term impact.

Our empirical results show that announcements affect trust and cooperation. The public trusts the organizations more if the organizations announce to send the transfer back. On average, organizations follow their announcement, but the rate of following is different for decision-makers in different tenures. Decision-makers in the last term in the organization do not follow a cooperative announcement as much as other decision-makers do. Although the effects of a cooperative announcement are short-lived at the decision-maker level, they are long-lasting at the committee level. With cooperative announcements, trust and cooperation rates remain high, even when all decision-makers have changed. Notice that one out of the two -re-specified message says that the committee will never cooperate, which does not make sense. As a result, the effects identified by our experiments can be considered to represent lower bounds of announcement effects in more natural set-ups.

We estimate what strategies the subjects play. While many studies have been conducted to analyze the indefinitely repeated Prisoners' Dilemma (for a recent overview, see, for example, Dal Bó and Fréchette, 2018), to the best of our knowledge, there are only two studies that experimentally examine indefinitely repeated trust games: Engle-Warnick and Slonim (2004, 2006). Their focus is on the comparison of finitely repeated and indefinitely repeated interactions of individual decision-makers. They do not consider the institutional design questions addressed in this paper. Dvorak and Fehrler (2019) study the effect of communication in the indefinitely repeated Prisoners' Dilemma and find, among other things, that pre-play communication is very effective in reducing strategic uncertainty and facilitates much higher cooperation rates. Relatedly, we find mission announcements tend to be conducive to trust in an organization.

Our study relates to the broad literature on social capital and trust between agents (Durlauf and Fafchamps, 2004; La Porta et al., 1997). Many empirical studies find that trust has benign economic consequences (Zak and Knack, 2001; Knack and Keefer, 1997; Guiso et al., 2004; Tabellini, 2010). Glaeser et al. (2000) focus on how to measure trust precisely and distinguish between trust and trustworthiness. The tension between long-lived organizations and relatively short-lived members is theoretically examined by Cremer (1986), Salant (1991) and Smith (1992). Xu and Potters (2018) and Offerman et al. (2001) study such set-ups in the laboratory.

In contrast to the present study, these studies consider cooperation between members of an organization rather than between an organization and an outside player (the public).

There are also studies that consider the trust of the public in a specific organization, in most cases, a central bank. Empirical studies find that macroeconomic outcomes influence trust in the European Central Bank (Fischer and Hahn, 2008; Gros and Roth, 2010; Farvaque and Mihailov, 2012; Ehrmann et al., 2013; Bursian and Fürth, 2015). The role of the institutional characteristics of central banks in mitigating the time-inconsistency problem in monetary policy, which goes back to Barro and Gordon (1983) and Kydland and Prescott (1977), has been considered by several theoretical analyses.

In particular, Sibert (2003) and Hansen and McMahon (2016) examine the benefits of collective and individual decision-making in central banks. Hansen and McMahon (2016) find differences in behaviors between new and old committee members, which is loosely related to a theoretical predictions we obtain for a collectivist organization with overlapping terms. In contrast to our approach, these studies examine signaling models with different types of central bankers. Other contributions to time-inconsistency problems and monetary policy committees include Dal Bo (2006), who examines optimal super-majority rules, and Riboni (2010), who focuses on the role of the status quo as the default option in committees. Kugler et al. (2007), who experimentally study trust between individuals and between groups, find that individuals are as trustworthy as groups. We extend their finding to a set-up where groups potentially enjoy the additional advantage that they interact with outsiders over longer time horizons compared to individuals.

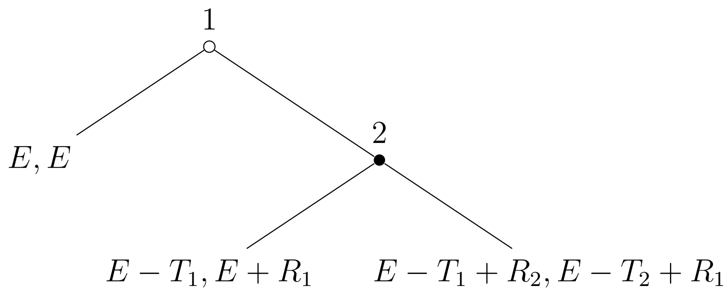
In the next section, we set up the model and derive theoretical predictions for the experiment. In Section 2.3, we describe our experimental design and briefly discuss the two pilot sessions. In Section 2.4, we present our main hypotheses and research questions. Section 2.5 presents the results of the study. The Appendix includes the proof of one of our theoretical predictions, details of our simulations for the power calculation, screen shots, and instructions of the experiment.

2.2 Model

2.2.1 Set-Up

We consider four versions of an indefinitely repeated trust game with binary choices. The first scenario, which considers an individualistic organization with a single decision-maker, is

Figure 2.1: Trust Game with Binary Choices (in Extensive and in Normal Form)



	$Send_2$	$Not\ Send_2$
$Send_1$	$E-T_1+R_2, E-T_2+R_1$	$E-T_1, E+R_1$
$Not\ Send_1$	E, E	E, E

labeled "I." The second scenario ("C") focuses on a collectivist organization with overlapping terms. Third, a collectivist organization with synchronized terms is described by model variant "CST." Finally, we consider a variant of the second scenario with the possible announcement of a mission ("CA").

We first describe the stage game for scenario I. The organization comprises three members: the decision-maker and two passive players. The decision-maker makes all decisions on behalf of the organization. The passive players make no choices in scenario I. The organization interacts with another player: the public. At the beginning of the stage, all players receive an endowment E. The public can choose to send T_1 ($T_1 < E$) to the organization. In this case, each of the three members receives R_1. We assume $3R_1 > T_1$ to capture efficiency gains from cooperation. If the public does not transfer resources to the decision-maker, the stage game ends and all players' payoffs are E.

If the organization receives a transfer, the decision-maker can choose whether the organization should keep the entire transfer, in which case all members' payoffs are $E + R_1$ and the payoff of the public is $E - T_1$, or whether to transfer a fixed sum T_2 ($T_2 < R_1$) per member to the public. If the organization makes transfers, public receives R_2. Thus the payoffs are $E_1 - T_1 + R_2$ for the public and $E + R_1 - T_2$ for members of the organization. Analogous to the assumption $3R_1 > T_1$, we impose $R_2 > 3T_2$ to describe efficiency gains from cooperation. We note that full cooperation, i.e. when both parties make transfers, maximizes aggregate payoffs. If no transfers are made, i.e. in the absence of cooperation, all players' payoffs are strictly lower in comparison.

It remains to describe how this stage game is extended to a repeated game in scenario I. The interaction between the public and the organization is of indefinite length: After every round there is a continuation probability $\delta \in (0, 1)$. With probability $(1-\delta)$ the interaction ends after every round.[6] Hence, the public has an indefinite time horizon, whereas all members of the organization have a fixed term in office, which we assume to be three periods. Terms overlap. In period $t = 1$, the decision-maker starts in the first period of her term. One passive member is in the second period of her term and one passive member is in the last period of her term.

Retiring members are always replaced by new members with fixed terms of three periods. Passive members are replaced by new passive members. The decision-maker is replaced by a new decision-maker. Members only receive payoffs while in office. New members always observe the entire history of choices made by all players.

Scenario C differs from scenario I only in that decisions are not made by a single individual but by all three members of the organization. Decisions on the transfers are taken by majority rule. As before, all members' payoffs are identical and terms overlap. The public observes the outcome of the vote but does not receive any information about individual votes.

Scenario CST, in turn, is almost identical to scenario C. The only difference is that the terms of the three decision-makers are synchronized. In period 1, all decision-makers are in their first term in office. After periods $t = 3, 6, 9, ...$, all decision-makers are replaced by new office holders.

For the last scenario, CA, we add an additional initial phase to scenario C. In this phase, the committee members who will be in charge in period 1 have the opportunity to vote on two available mission statements regarding the future choices of the organization. For simplicity, we allow only for the statements that "send will be chosen" and that "send will not be chosen." The statement selected by the majority will be revealed to the public and all future members of the organization. The exact pattern of votes will not be revealed.

[6]Note that this is equivalent to assuming that time $t = 1, 2, ...$ is infinite and payoffs in future periods are discounted by the common factor $\delta \in (0, 1)$.

The equilibrium concept is subgame-perfect Nash equilibrium. As is well-known, voting games often have multiple equilibria as voters are indifferent when they are never pivotal. Thus we impose the following restriction: If a decision-maker is not pivotal with certainty, she votes for the option that would maximize her payoffs.[7]

This restriction will allow us to identify a unique equilibrium in scenario CST. However, scenario C admits multiple equilibria even if this refinement is imposed. It will be of particular interest whether subjects in the experiment will be able to coordinate on an equilibrium with full cooperation, i.e. one that maximizes aggregate payoffs.

2.2.2 Theoretical Predictions

In the following, we derive the theoretical predictions for the different scenarios.

An individualistic organization (I) Applying a standard backward-induction argument leads to the finding that there is a unique subgame-perfect equilibrium, in which no transfers occur. We summarize this result as follows:

Theoretical Prediction 1. *An individualistic organization leads to no cooperation. In particular, the public never sends transfers to the organization. Decision-makers would never send transfers to the public.*

A collectivist organization with synchronized terms (CST) The equilibria are similar to the ones in scenario I. Consider, e.g., period 3, after which all decision-makers leave. If a decision-maker is pivotal with positive probability, it clearly maximizes her payoffs to send no transfers to the public. If she is pivotal with probability zero, her action cannot affect the outcome of the vote. Because of our equilibrium refinement, she nevertheless votes against sending transfers to the public. As a consequence, conditional on having received a transfer, all decision-makers vote in favor of not sending a transfer to the public. Anticipating this behavior, the public does not send transfers in the third period. Backward induction yields that also in periods 2 and 1, no player makes transfers. The theoretical prediction for the CST scenario is thus identical to the prediction for scenario I:

[7]Obviously, this restriction has no bite in scenario I.

Theoretical Prediction 2. *A collectivist organization with synchronized terms leads to no cooperation. In particular, the public never sends transfers to the organization. Decision-makers would never vote in favor of sending transfers to the public.*

We would like to highlight that this prediction relies on our additional refinement about the behavior of non-pivotal players. Without this refinement, equilibria with cooperation would exist. Suppose all decision-makers always voted in favor of sending transfers to the public. Then no profitable deviation would exist for individual members of the organization. As a consequence, it would be optimal for the public to send transfers to the organization.

A collectivist organization with overlapping terms (C) As a next step, we consider a collectivist organization with overlapping terms. In every period, there is a majority of decision-makers who remain in office for at least one additional period and therefore potentially benefit from future cooperation. Due to this feature, cooperation can occur in equilibrium.

Theoretical Prediction 3. *In a collectivist organization with overlapping terms, multiple equilibria exist. In particular, the following behaviors may occur in equilibrium:*

1. *The public never sends transfers to the organization, i.e. there is no cooperation.*
2. *Suppose that*
$$T_2 \leq \frac{\delta}{1+\delta} R_1. \tag{2.1}$$
Then the following behaviors are supported by an equilibrium: The public always sends transfers to the organization. The organization always sends transfers in return. Hence there is full cooperation and aggregate payoffs are maximal in every period.

Proof. See Appendix B.1. □

Thus Condition (2.1) ensures that cooperation is possible. Intuitively, it guarantees that the costs of cooperating incurred by members of the organization, which are associated with T_2, are sufficiently small such that decision-makers do not wish to forgo the future gains from cooperation, which are positively influenced by R_1. As explained in Appendix B.1, cooperation cannot occur if (2.1) is violated.

A collectivist organization with the announcement of a mission statement (CA) As scenario C involves multiple equilibria, it is possible to construct equilibria with mission statements where the announcement has an effect on future behavior despite the fact that the statements represent cheap talk. While no clear-cut theoretical prediction emerges, the following results can be plausibly expected. First, we would expect the organization to announce that it intends to achieve cooperation, i.e. that it will transfer funds to the public. Second, it appears possible that such an announcement will help the players to select the payoff-maximizing equilibrium with full cooperation.

2.3 Experimental Design

We propose to implement the indefinitely repeated binary trust game in four treatments in a between-subject design. In all treatments, the public repeatedly interacted with the organization. For all supergames, the continuation probability after every round δ is 5/6.[8] Thus, the expected length of each supergame is six. We implement indefinite repetition through random termination. To keep the length of the supergames constant across treatments, we generate three sequences beforehand and use them to determine the length of each supergame. There are nine, five and nine supergames with 41, 40 and 39 rounds in total for the three sequences respectively and all three sequences will be implemented for one-third of the subjects of each treatment.[9] The treatments vary in the organizational structure of the organization.[10]

- **I** (Treatment with an individualistic organizational structure.) One individual decision-maker interacts repeatedly for 3 periods with the public. At the end of the third period, the current decision-maker is replaced by a new decision-maker.

- **C** (Treatment with a collectivist organizational structure and overlapping terms.) The committee consists of three decision-makers of overlapping generations. At the end of

[8] A 'supergame' is one indefinitely repeated game, including all rounds until its random termination.

[9] We used Stata to generate three sequences of uniformly distributed random numbers between 0 and 1 with seeds 7, 8 and 9. We used seeds 1–6 in our previous research projects; therefore, we started from 7. The three sequences are denoted as $\{r_n\}_i = \{r_1, r_2, ...\}_i$, where $i = 7, 8, 9$ indicates the seed underlying the sequence, and $n \in \mathbb{N}$. The first supergame has n_1 rounds if $r_{n_1} \leq 1/6$ and for all $n < n_1$, $r_n > 1/6$. The second supergame has $n_2 - n_1$ rounds if $r_{n_2} \leq 1/6$; for all $n_1 < n < n_2$, $r_n > 1/6$. The procedure is repeated to find the lengths of all supergames. The resulting length sequences are SQ1 $(3, 2, 1, 17, 1, 1, 7, 7, 2)$, SQ2 $(16, 10, 2, 11, 1)$ and SQ3 $(6, 2, 2, 15, 2, 1, 5, 3, 3)$.

[10] See Appendix B.5 for screenshots and instructions.

every period, the subject representing the oldest generation exits the committee. At the beginning of every period, a subject enters the committee as the young member.

CST (Treatment with a collectivist organizational structure and synchronized terms.) Three decision-makers interact repeatedly for 3 periods with the public. At the end of the third period, all members retire an are replaced by three new members.

CA (Treatment with a collectivist organizational structure, overlapping terms and mission announcement.) The treatment differs from C with an additional mission statement stage. Before each supergame, the committee votes on one of the two announcements: "We will always play 'send'." and "We will always play 'not send'." The committee announcement is chosen by the majority rule and remains displayed on the screens of all subjects in the same group for the entire supergame.[11]

In all treatments, the public consists of a single subject who stays in position until the supergame ends. In each period, we let subjects play the normal-form binary trust game, that is, the public and decision-makers decide on actions simultaneously rather than sequentially. In the treatments C, CST and CA, the decision of the organization (that we call committee in the instructions) is made by simple majority rule.

At the end of each round in all treatments, the public is informed of the individual decision-maker's decision or committee decision if and only if the public chooses *Send*. The individual votes of committee members in C, CA and CST remain secret to the public. The individual decision-maker always receives feedback on the decision of the public. Committee members receive feedback on individual votes, committee decision and the public's choice.

Matching Groups For treatment I, there are 27 subjects per session. Each session consists of three matching groups. Each matching group is further divided into three groups. Each group consists of one subject in the role of the public, an individual decision-maker, and a waiting subject. Sessions start by randomly assigning subjects into the three matching groups. Within each matching group, we assign them then randomly into three groups and into the role of public, decision-maker or into the waiting pool.[12] Subjects assigned to the role of the

[11] As mentioned in the Introduction, we will only implement this treatment if the trust rate is not higher than 85% in treatment C. In that case, there would not be enough scope for improvement.

[12] The role of the waiting pool will be explained in detail below.

public keep this role throughout the whole session. Before the start of each supergame, we randomly rematch subjects within their matching group. The subjects that are not in the role of the public are again randomly assigned to a role. For all sessions, each matching group is assigned a different sequence. Hence, the lengths and numbers of supergames differ among the three matching groups.

Recall that in the theoretical model, a decision-maker serves a finite number of terms. Subsequently, the decision maker retires. Implementing this setup one-to-one in the laboratory would be difficult because it would require (indefinitely) many subjects. To solve this problem, we must allow for re-entry while keeping the chance of re-entry into the same group sufficiently low to avoid repeated game effects in the last term of a decision-maker. We chose the following implementation: the individual decision-maker interacts repeatedly with the public in the same group for three periods. Afterwards, she is replaced by the waiting subject of that group. She then waits in the next group for three periods before becoming a decision-maker there. For example, a decision-maker who retires from group 1 continues by waiting in group 2 for three periods before becoming a decision-maker there. Similarly, the next step for a decision-maker who retires from group 3 is to wait in group 1 for three periods before becoming a decision-maker in that group. Hence, it takes 13 periods for a retired decision-maker to re-enter her initial group (and at first in the waiting pool) and 16 periods to become a decision-maker again in her initial group. With our continuation probability $\delta = 5/6$, the chance of the latter is only approximately 5%, which is sufficiently low to not affect our theoretical predictions.

For treatments C, CST and CA, each session consists of one matching group. Each matching group again consists of three groups, and each group consists of one subject in the role of the public, three decision-makers in the committee of either overlapping or synchronized terms, and a waiting pool of three subjects.

Analogous to treatment I, subjects can re-enter, but the chance is equally small. All retired decision-makers move into the next group, and spend three periods waiting there before entering that group's committee. Hence, it again takes 13 periods for a retired decision-maker to re-enter her initial group and 16 periods to become a decision-maker again in that group. First, this design ensures that the lab implementation is close to the theoretical scenario, and second, that the treatments are identical with respect to the low re-entry probability.

In all treatments, the choice history of both the decision-makers and the public is visible to all waiting subjects in the same group.

Our simulations (see Appendix B.2) suggest that we have enough power ($> 87\%$) to detect effect sizes of 15 percentage points at the 5% level with a one-sided $t-$test and 6 matching groups per treatment.[13]

Payment A potential concern regarding the individualistic setup I could be that the public's choice only influences the payment of one decision-maker, whereas it affects three players in the other treatments. Therefore, we design the payment in treatment I in such a way that the choice of the public also influences three players. To do so, we randomly draw two subjects from another matching group as "passive members" of the organization. They are paid the same amount as the decision-maker, but they do not engage in decision-making. Subjects are not aware of whether they are chosen as passive members until the end of the session.

In all sessions, subjects are payed a show-up fee of EUR 5 plus their accumulated earnings over all rounds. The points they earn are exchanged for euros at an exchange rate of 5 cents per point. As subjects who are not in the role of the public are in the waiting pool in half of the rounds (in expectation), they would earn substantially less than the subjects in the role of the public in the C, CST and CA treatments. In the I treatment, they earn enough extra points from being picked as passive assistants. To raise the average earnings of the subjects that are not in the role of the public in C, CST, and CA, we pay waiting subjects a fixed wage of five points per round that they have to wait.

Experimental Parameters Figure 2 shows the stage-game payoffs for all treatments. The public is the row player and the decision-makers are the column players. Players are endowed with $E = 5$. If the public does not transfer, players end up with their endowments. If the public transfers $T_1 = 4$, the (three) committee members receive $R_1 = 6$; that is, the transferred amount is multiplied by a factor of (3*)1.5. When a transfer is received, the decision-maker(s) can send back $T_2 = 2$, which reduces her payoff to 9. The public receives $R_2 = 8$, that is, the back transfer (from all three committee members) T_2 is multiplied by a factor of four (over three).

We choose these parameters for the following reasons. First, they ensure that Condition 1 (in Theoretical Prediction 3) holds; i.e. in collectivist organization with overlapping terms,

[13]For the simulations we assumed trust probabilities of 0.4 for I (CST or C) and 0.55% for C (C or CA). Power would increase substantially if we assumed trust probabilities closer to 0 or 1 (for example, 0.1 and 0.25 or 0.8 and 0.95). This is because the variance of the Bernoulli distribution, $p(1-p)$, is largest at $p = 0.5$.

cooperation between the public and decision-makers can be sustained with $\delta = 5/6$. Second, the public is indifferent with regard to 'Send' and 'Not Send' (in the absence of repeated game effects) if they consider decision-makers to be trustworthy with a probability of 50%. Third, both parties receive the same payoffs if 'Send' is chosen by both and if 'Not Send' is chosen by the public. Finally, the cooperation payoff 9 is efficient. [14]

Figure 2.2: Stage Game Parameters

	$Send_2$	$Not\ Send_2$
$Send_1$	$9, 9$	$1, 11$
$Not\ Send_1$	$5, 5$	$5, 5$

Sessions All treatments are programmed in z-Tree (Fischbacher, 2007). Subjects are recruited via hroot (Bock et al., 2014). We run 19 sessions in total (3 * 6 sessions with one matching group of 21 subjects each for the C, the CA and CST treatments, and one sessions with three matching groups of nine subjects each for the I treatment) with a total of 405 subjects.[15] We conduct 10 sessions in WISO Experimental Lab of the University of Hamburg in 2021 and 2022, and 9 in LakeLab of the University of Konstanz in 2022 and 2023. [16] Each session takes less than two hours and subjects were able to understand the matching protocol. Subjects in Lakelab additionally answer seven extra questions at the end of the experiment in the questionnaire. In the questions, subjects are asked to look back on their decision-making and answer questions such as how important they thought their decisions are and whether organizational structures have an impact on their decisions. The complete questionnaire and a summary of the answers to the questionnaire can be found in Appendix B.6. No treatment difference in the answers is detected between the treatments.

[14]Reciprocating transfers induces a loss of 2 for the decision-maker in individualistic treatment, and an aggregate loss of 6 in collectivist treatments. The public gains 8. Hence, cooperation increases efficiency.

[15]We intend to run 20 sessions in total, with 3 * 6 sessions with one matching group of 21 subjects each for C, CST and the CA treatments, and two sessions with three matching groups of nine subjects each for the I treatment. These numbers stem from our power calculations. The number of matching groups provides sufficient power (> 87%) to detect effect sizes of 15 percentage points at the 5% level using a one-sided t-test. However, we had difficulty filling up our sessions. By the time this paper is being written, there is still one session yet to be conducted.

[16]We divide our sessions so that each lab conducts half of the sessions with all treatments of every random sequence.

2.4 Main Hypotheses and Research Questions

For the theoretical reasons outlined in the previous sections and judging by results from previous studies comparing finite and indefinite repetitions of a stage game, we expect higher cooperation in the C treatment than in the CST and I treatments. Moreover, we expect that an announcement of an organizational mission leads to higher trust in the CA treatment than in the C treatment. As is usual in the analysis of experiments with indefinitely repeated games, we will give subjects time to learn and focus our attention on the last seven supergames of sequences 1 and 3 and the last three supergames of sequence 2 to test our hypotheses.[17]

Based on these observations we formulate our hypotheses, which directly follow from the theoretical predictions in Section 3.

H1s: *Cooperation rates will be higher in the C treatment than in each of the CST and I treatments in the last five supergames. Cooperation rates will also be higher in the CA treatment than in the C treatment in the last five supergames.*

We test the three corresponding H0s by comparing the cooperation rates within the pairs of treatments C–I, C–CST, and CA–C. We treat each group as an independent observation and cluster at the matching group level for one-sided $t-$tests on the equality of the cooperation rates. We further run a two-sided $t-$test of the null hypothesis that there is no difference in the trust rate between I and CST, which we predict is the case. Then, we split the observations in the CA treatment based on the announcement results. Although no clear theoretical predictions could be made on the frequency of cooperation between groups that announce to send back and groups that announce not to, we expect that committee members will follow their announcement and that the public will trust committees that announces to be cooperative. We compare the cooperation rates between groups of announcements using a one-sided $t-$test.

Apart from our main hypotheses, we answer the following research questions:

[17]We exclude 2 supergames of each sequence from the data set by treating them as learning phases. Since there are 9 supergames for sequence 1 and 3, 5 supergames for sequence, thus 7 supergames left for sequence 1 and 3, while 3 left for sequence 2.

Question 1: Does decision-makers' tenure affect their choices?

Our stage-game parameters theoretically ensure that in treatment C, the decision-makers in the final term have no incentive to send back the transfer, whereas those in the first and second terms have an incentive to do so in the equilibrium with cooperation that we consider. Thus, we expect that the frequency of sending back is lower for decision-makers in the final term. For treatment CST, based on our second theoretical prediction that cooperation is not possible, decision makers vote "Not Send" in each term, so their tenure does not affect their decisions. For treatment I, by backward induction, cooperation is not possible; however, since the decision-makers stay in office for three terms, it is beneficial for them to send back the transfer in the first two terms and keep the transfer only in the last term. This repeated game effect leads to a low frequency of sending back in term 3. Finally, it is unclear how a simple announcement mechanism interacts with tenure. If decision-makers adhere to group announcements, we expect to see less difference among terms as they vote for the same decision in every term.

Question 2: How does cooperation change over rounds?

In a collectivist organization with overlapping terms, the composition of the committee remains stable in the sense that there is always one member who has just joined, one member who joined one period ago, and one member who is about to leave. Therefore the frequency of playing "Send" on the committee level and the overall cooperation level should be stable over time. An individualistic organization does not have this feature because each decision-maker interacts with the public for three periods: the public's trust will decrease in round 3 if the public expects the decision maker to not send back, which will lead to a reduction in cooperation rates in round 3.

Question 3: Which strategies do subjects play?

Strategy choices have been studied extensively for the repeated Prisoner's Dilemma (for an overview of the findings, see for example, Dal Bó and Fréchette, 2018). By contrast, with the exception of Engle-Warnick and Slonim (2004, 2006), there is no research on strategy choices in repeated trust games. To fill this gap for our repeated trust game between an organization and the public, we build on the strategy frequency estimation method (SFEM) introduced by Dal Bó and Fréchette (2011) and use the R package stratEst, which was developed by Dvorak (2019) and first used in Dvorak and Fehrler (2019). The SFEM is frequently used to obtain maximum-likelihood estimates of the shares of a candidate set of strategies in experimental data. We base our estimation on a candidate set of strategies from Engle-Warnick and Slonim (2006) with 18 pure strategies for the trustors. We add one additional pure strategy according to which the public sends as long as the committee sends as well, otherwise plays “Not Send” until the end of round 3 and round 6, then returns to “Send”. This strategy is a potential punishment strategy for treatment I and CST.

2.5 Results

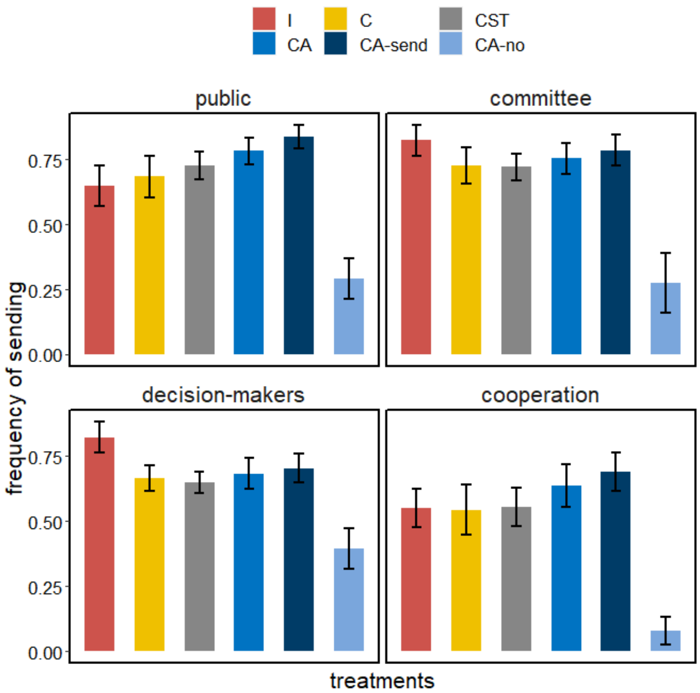
To account for learning, we exclude the first two supergames. All the results in this section are based on data from the last seven supergames for random sequences 1 and 3, or the last three supergames for random sequence 2. We begin by testing our main hypotheses.

2.5.1 Frequency of Sending

As the binary trust game is asymmetric, we look at the frequency that the public and decision-makers within a committee play “Send” separately. The first column of Figure 2.3 shows the results of the average frequency of sending for the public and decision-makers inside a committee. The second column corresponds to the committee’s mean rate of sending and their mean rate of cooperation. Cooperates rates are defined as the rates of mutual sending. We treat each committee (group) as an independent observation to calculate the average and cluster at the matching group level to obtain clustered standard errors. The first four bars in each subplot represent the four treatments I, C, CST, and CA. We further split the CA treatment into two scenarios, “CA-send” and “CA-no”, based on whether the committee promises to play “Send” in the announcement stage. Table 2.1 summarizes the mean differences and $p-$values between the treatments for each plot.

For the public, the average rate of sending is a measure of average trust in the committee. A higher sending rate is interpreted as higher trust. Overall, the trust rate is high across all the treatments. The mean rate of trust is the same between treatments. The public transfers funds more often when the committee announces to send back than when they do not make an announcement and when they promise not to. For the decision-makers, the frequency of sending is higher in treatment I than in treatment C and is different between I and CST. There is no difference between treatment CST and C, and neither is there difference between treatment C

Figure 2.3: Plot of the average frequency of sending



Note: This figure plots the average frequency of sending of the public, committee, the decision-makers of the committee, and the cooperation rates between the public and the committee for each treatment. Treatment CA is further split into CA-send and CA-no, depending on whether the committee announces sending back the transfer. The average is calculated at the group level and clustered at the matching group level. Error bars indicate clustered standard deviation.

and CA. When the committee promises to play "Not Send," the frequency of sending is both lower than when the committee does not promise and when the committee promises to play "Send".

We look further at the mean frequency of sending at the committee level. The first row of the second column in Figure 2.3 shows the results. The mean differences and $p-$values are summarized in Table (3) of 2.1. For treatment I, because the committee consists of an individualistic decision-maker, the mean rate of sending of the committee is the same as the mean rate of sending of the decision-makers. There is no difference in the mean across the four treatments. The mean when the committee announces to play "Not Send" is lower than both when no announcement is made and when the announcement is "Send".

Finally, we define cooperation as mutual sending. The second row of the second column in Figure 2.3 shows the mean frequency of cooperation. The mean differences and $p-$values are summarized in Table 2.2. The means are not statistically different between the treatments. Announcing to play "Not Send" decreases the mean cooperation.

We cannot reject our null hypotheses because the means of sending and cooperating are not statistically different. We find empirical evidence confirming the importance of announcements. Announcing to play "Send" increases the average trust of the public. On average, the committee adheres to their announcements. Announcing to play "Not Send" decreases not only public trust but also the frequency of sending on the committee's side and the mean rate of cooperation.

Result 1: Individualist and collectivist organizational structures do not affect the mean trust and cooperation rates between I and C, CST, and C. Overall, the frequency of trust is high in all treatments, although sending is risky for the public.

Result 2: Most committees announce sending back the transfer ($M = 0.89$, $SD = 0.03$). Announcements affect trust and cooperation. The public trusts the committee more if it announces sending the transfer back. The committee follows their announcement. The rate of sending on the committee's side is low when they promise not to send back the transfer. Announcing to play "Not Send" harms cooperation.

2.5.2 The Effect of Tenure

We next compare the decisions of decision-makers of the same tenure across treatments and the decisions of decision-makers of different tenures within the same treatment. The results are shown in Figure 2.4. For simplicity, the decision-makers in the first, second, and last terms are abbreviated as M1, M2, and M3, respectively.

We again split treatment CA into CA-send and CA-no based on whether "Send" has been announced. The mean frequency that M1 announces to send is 0.87 ($SD = 0.03$). For M2, the mean frequency is 0.73 ($SD = 0.06$). And it is 0.75 ($SD = 0.03$) for M3.

For the same tenure between treatments, we find that M1 sends back more often in treatment I than in C, and the mean is different between I and CST (I vs. C: $p_{M1} = .025$; CST vs. C: $p_{M1} = .191$; C vs. CA: $p_{M1} = .382$; I vs. CST: $p_{M1} = .002$). When the committee announces to play "Not send", the mean of sending decreases for M1 (CA vs. CA-send: $p_{M1} = .496$; CA-no vs. CA: $p_{M1} = .003$; CA-no vs. CA-send: $p_{M1} = .011$). Similar results are observed for M2.

Table 2.1: Mean differences and $p-$values for treatment comparisons

treatments		mean difference	p-value
I	C	-0.04	.360
	CST	-0.08	.355
CST	C	0.04	.325
C	CA	-0.10	.141
	CA-send	-0.15	.043
CA-no	C	-0.39	$< .001$
	CA-send	-0.55	$< .001$

(1) public

treatments		mean difference	p-value
I	C	0.16	.015
	CST	0.17	.008
CST	C	-0.02	.395
C	CA	-0.02	.408
	CA-send	-0.04	.298
CA-no	C	-0.27	.001
	CA-send	-0.31	.006

(2) committee

treatments		mean difference	p-value
I	C	0.10	.130
	CST	0.10	.159
CST	C	< 0.01	.475
C	CA	-0.03	.376
	CA-send	-0.06	.251
CA-no	C	-0.45	$< .001$
	CA-send	-0.51	$< .001$

(3) decision-makers

Note: The tables summarize the mean differences and $p-$value between treatments. From top to bottom, the tables are based the choices of the public, the committee, and the decision-makers.

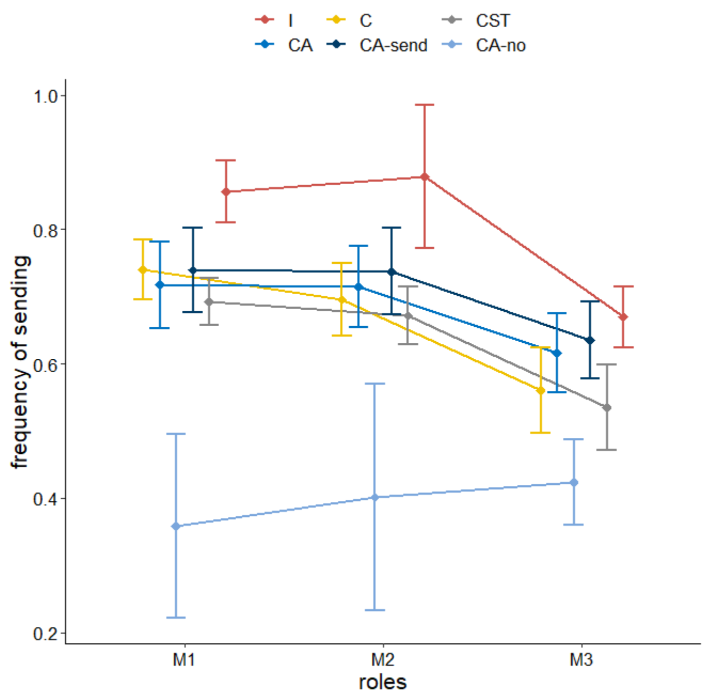
Table 2.2: Mean differences and $p-$values for treatment comparisons

treatments		mean difference	p-value
I	C	< 0.01	.476
	CST	< 0.01	.962
CST	C	0.01	.461
C	CA	-0.09	.218
	CA-send	-0.15	.105
CA-no	C	-0.46	$< .001$
	CA-send	-0.61	$< .001$

(4) cooperation rates

Note: This table summarize the mean differences in cooperation rates and the corresponding $p-$value between treatments.

Figure 2.4: Plot of average frequency of sending of decision-makers in term 1, 2 and 3 in committee



Note: This figure shows the average frequency of sending of the decision-makers in the first, second and third term in the committee. Treatment CA is further split into CA-send and CA-no depending on whether "Send" has been announced. The average is at committee level and we cluster on the matching groups for standard deviation. Error bars show the clustered standard deviations.

Decision-makers in their second term send back more often in treatment I than in C, and the rate of sending back differs between I and CST (I vs. C: $p_{M2} = .042$; CST vs. C: $p_{M2} = .356$; C vs. CA: $p_{M2} = .403$; I vs. CST: $p_{M2} = .041$). Announcing to play "Not Send" by the committee reduces the mean compared with no announcement (CA vs. CA-send: $p_{M2} = .303$; CA-no vs. CA: $p_{M2} = .040$; CA-no vs. CA-send: $p_{M2} = .064$). For M3, the mean is not statistically different between the four treatments (I vs. C: $p_{M3} = .067$; CST vs. C: $p_{M3} = .381$; C vs. CA: $p_{M3} = .249$; I vs. CST: $p_{M3} = .063$). A "Not Send" announcement reduces the mean compared to a "Send" announcement (CA vs. CA-send: $p_{M3} = .178$; CA-no vs. CA: $p_{M3} = .054$; CA-no vs. CA-send: $p_{M3} = .018$).

Within each treatment, we compare the effect of being in different terms on the mean rate of sending by regressing the average frequency of sending on the term in office. We treat M1 as the base group. Table 2.3 shows the results for each treatment, as well as for CA-send and CA-no. For treatment I and treatment CST, being in the final term in committee decreases the mean rate of sending. For treatment C, although theory predicts that both M1 and M2 should send back because of the future benefit of cooperation, the mean frequency of sending decreases as the tenure increases. Announcement however, changes this pattern. The mean sending rate does not decrease from M1 to M2 if announcement is possible, regardless of what the committee has promised. Although announcements help M2 to play "Send" more often, this effect is limited and does not extend to M3. Even when the committee has promised to "Send", the mean frequency that M3 plays "Send" is still decreased.

Our results above show that announcements affect the decisions of decision-makers of all tenures, especially when "Not Send" has been announced. Our results further indicate that there is a last term effect in treatment I, CST, and CA that the decision-makers play "Send" in the first two periods, but keep the funds in the last period, which makes the mean rate of sending higher for M1 and M2 than M3. However, according to Figure 2.4, the average sending rate for M3 is still quite high. This may imply that a high share of decision-makers do not reciprocate for future gains, but are reciprocators who reciprocate even with the absence of future gains. Finally, according to our stage game parameters, both M1 and M2 in treatment C have an incentive to send. Empirically, we find that M2 sends less often than M1. This difference in the sending rates between M1 and M2 decreases when announcements are allowed. This effect of announcement is short-lived and does not carry over to M3. M3 leaves the committee in the next round, and there is no incentive for them to play "Send", so they decrease the rate of sending in the last term.

Table 2.3: Effects of tenure on mean rate of sending

	I	C	CST	CA	CA-send	CA-no
constant	0.86***	0.74***	0.69***	0.72***	0.73***	0.36***
	(0.05)	(0.05)	(0.03)	(0.07)	(0.06)	(0.12)
M2	0.02	-0.04***	-0.02	<-0.01	<-0.01	0.04
	(0.08)	(0.02)	(0.05)	(0.01)	(0.03)	(0.17)
M3	-0.19***	-0.18***	-0.15***	-0.10**	-0.11*	0.07
	(0.06)	(0.04)	(0.05)	(0.04)	(0.05)	(0.17)
N	27	54	54	54	54	27
adjusted-R^2	0.19	0.20	0.18	0.04	0.05	-0.08

Note: This table shows the results of linear regressions. The dependent variable is the mean rate of sending at the committee level, and the explanatory variable is the term, with M1 as the base group. Standard errors in parentheses are clustered at the matching group level. $***$ ($**$,$*$) indicates the significance at the 1 (5,10)% level.

Result 3: The mean frequency of playing “Send” decreases over tenure in treatment C, which contradicts our theoretical prediction. The announcement helps reduce the difference in sending between M1 and M2, but this effect is limited and does not carry over to M3.

Result 4: There is a last term effect in treatment I, CST and C. Decision-makers in these treatments send often in term 1 and term 2, but keep the funds in the last term.

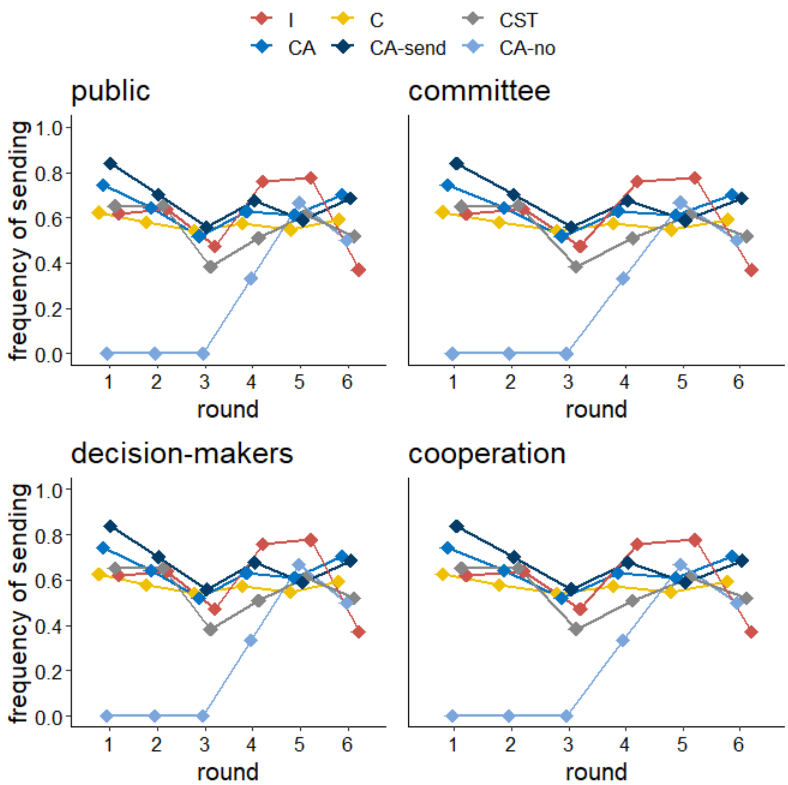
The above results do not tell us how the sending rates change over the rounds for a collectivist organizational structure with overlapping generations. We look at over-the-round changes in the next subsection and focus not only on the average sending rate of decision-makers inside a committee but also on the public and cooperation rates.

2.5.3 Changes Over Rounds

Figure 2.5 shows the changes in the average frequency of sending and cooperation over the rounds. The four panels correspond to the choices of the public, committee, decision-makers and the cooperation rates. The average is calculated at the committee level over all supergames for every round. We examine the first six rounds because the continuation probability of our indefinitely repeated trust game is 5/6. The expected length of each supergame is six.

According to Figure 2.5, the frequency of sending for the committee, decision-makers, and cooperation rates decreases in rounds 3 and 6 for treatment I. For treatment C and CA, the changes in the mean rate of sending and cooperation are much smaller than those in treatments I and CST. Announcing to play "Send" stabilizes the mean rate of sending and cooperating over rounds. However, if a "Not Send" announcement is made, the mean rate of sending and cooperation is low for the first three periods, but increases in round four when all old decision-makers have left. The new committee members do not follow the announcement made by the old committee.

Figure 2.5: Plot of average frequency of sending of decision-makers over rounds



Note: This figure shows the changes in the mean frequency of cooperation over the rounds for all treatments. Treatment CA is further split into CA-send and CA-no depending on whether or not "Send" has been announced. Panel A shows the public's decisions. Panel B shows the committee's voting decisions. Panel C shows the average decisions of the committee members. Panel D shows the average cooperation rates over the rounds. The average is on calculated on committee level.

To test whether a repeated game effect exists and to examine how it interacts with different treatments, we define a dummy variable which takes on value 1 if the round is a multiple

Table 2.4: The round effect on the mean rate of sending and cooperation

	public		committee		decision-makers		cooperation	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
constant	0.72***	0.72***	0.73***	0.73***	0.67***	0.67***	0.58***	0.58***
	(0.07)	(0.07)	(0.08)	(0.08)	(0.06)	(0.06)	(0.10)	(0.10)
MULT3	-0.06	-0.06	-0.02	-0.02	<0.01	<0.01	-0.01	-0.01
	(0.06)	(0.06)	(0.02)	(0.02)	(<0.01)	(<0.01)	(0.04)	(0.04)
I	0.04	0.04	0.13	0.13	0.19**	0.19**	0.12	0.12
	(0.11)	(0.11)	(0.11)	(0.11)	(0.09)	(0.09)	(0.13)	(0.13)
CST	0.03	0.03	0.03	0.03	<0.01	<0.01	0.03	0.03
	(0.09)	(0.09)	(0.10)	(0.10)	(0.07)	(0.07)	(0.12)	(0.12)
CA	0.10	-	0.02	-	0.01	-	0.08	-
	(0.08)		(0.11)		(0.08)		(0.13)	
CA-send	-	0.14	-	0.05	-	0.03	-	0.12
		(0.09)		(0.11)		(0.08)		(0.13)
CA-no	-	-0.38***	-	-0.38***	-	-0.25***	-	-0.44***
		(0.09)		(0.14)		(0.10)		(0.13)
MULT3×I	-0.11	-0.11	-0.18***	-0.18***	-0.20***	-0.20***	-0.26***	-0.26***
	(0.16)	(0.16)	(0.06)	(0.06)	(0.06)	(0.06)	(0.09)	(0.09)
MULT3×CST	0.02	0.02	-0.12	-0.12	-0.11*	-0.11*	-0.14**	-0.14**
	(0.08)	(0.08)	(0.10)	(0.10)	(0.06)	(0.06)	(0.06)	(0.06)
MULT3×CA	-0.02	-	0.04	-	0.02	-	-0.03	-
	(0.07)		(0.05)		(0.04)		(0.05)	
MULT3×CA-send	-	-0.04	-	<0.01	-	<-0.01	-	-0.07
		(0.08)		(0.05)		(0.04)		(0.06)
MULT3×CA-no	-	0.14	-	0.08	-	0.05	-	0.04
		(0.19)		(0.18)		(0.11)		(0.08)
N	378	401	378	401	378	401	378	401
adjusted-R^2	0.01	0.09	0.01	0.08	0.06	0.11	0.02	0.11

Note: The table presents the regression results. The independent variable for the first six columns is the mean rate of sending of the public, the committee and the decision-makers of each committee in every round over all supergames. The independent variable for the last two columns is the mean rate of cooperation at the committee level for all supergames in every round. The explanatory variables are 1) MULT3, a dummy variable that takes on value 1 if the round is a multiple of 3; 2) treatment categories; and 3) the interaction of MULT3 and the treatments. The base group is treatment C. Column (2) splits CA into CA-send and CA-no, whereas column (1) does not. Numbers in parentheses are standard deviations that are two-way clustered at both the committee and matching group levels. $***$ ($**$,$*$) indicates the significance at the 1 (5,10)% level.

of 3. We regress the average frequency of sending and cooperation on this dummy, on the treatment category that treats treatment C as the base, and on the interaction of the round dummy and treatment category. Table 2.4 summarizes the results. For the first six columns, the independent variable is the mean rate of sending of the public, committee, or decision-makers of each committee in every round over all supergames. The independent variable for the last two columns is the mean rate of cooperation on the committee level over all supergames in every round. The standard errors in parentheses are two-way clustered at both the committee and matching group levels. Column (1) uses all data from the treatment CA. Column (2) splits CA into CA-send and CA-no.

According to Table 2.4, the public does not reduce the frequency of sending in rounds 3 and 6 compared with treatment C. The committee sends less often in rounds 3 and 6 in treatment I than in treatment C. When looking at the decision-makers, those in treatment I and CST reduce the frequency of sending when they are in the last term. Subjects from treatment I and CST cooperate less often than those from C in rounds 3 and 6. Our regression results confirm that there is a last-round effect in treatment I and CST, but not in C and CA. The last round effect is only observed for committee members but not for the public. A collectivist organizational structure with overlapping terms helps stabilize the rates of sending and cooperation over rounds. It is worth noting that although the announcement effects are short-lived, as the effect does not carry over to M3, they are still there in the fourth round, where all decision-makers have changed. The mean rates of sending and cooperation are both high in round 4-6 when a cooperative announcement has been made.

Result 5: The trust of the public in the committee is unaffected by the last round. The last-round effect exists only for the committee and the decision-makers inside the committee.

Result 6: Although a collectivist organization with overlapping terms does not increase the rate of sending and cooperation, it helps to keep them stable over rounds.

Result 7: The effect of announcement on the choices of the committee members depend on the tenure of the members, but it has a long-lasting time effect. Although the effect of a cooperative announcement is short-lived, as it does not carry over to M3, this effect is still present in round 4 when the committee has changed.

2.5.4 Choices After Memory-one Histories

We estimate the probability of playing "Send" after memory-one histories. We focus on the voting decisions of the committee instead of on the subject level.[18] We calculate the probability of playing "Send" after every memory-one history for each supergame. The results are summarized in Table 2.5.

The public and the committee condition on different memory-one histories. Denote σ as the probability of playing "Send". This probability conditions on one of the four possible memory-one histories ($\emptyset$, ss, sn, n). $\emptyset$ indicates the first round with no history. ss is when both the public and the committee have played "Send" in the previous round. sn is when the public has played "Send" but the committee has not. n is the case in which the public has chosen not to send the fund in the previous round. In this case, they could not see what the committee has chosen in that round. The corresponding cooperation rates after these histories are ($\sigma_{\emptyset}$, σ_{ss}, σ_{sn}, σ_n). We estimate the probabilities and the results are shown in Table (1).

We assume that the choices of the committee in the previous round affects the choices of the committee in the current round. Moreover, the choice of the committee is affected by the public. Thus, the memory-one histories are ($\emptyset$, ss, sn, ns, nn). The first letter in the history represents the public's choice in the previous round. The second letter is the committee's decision in the previous round. The estimated results for the committee are in Table (2).

In treatment CA, the decisions of both the public and the committee are further affected by the announcement. Taking announcement into account, the memory-one histories for the public become ($\emptyset$Send, ssSend, snSend, nSend, $\emptyset$No, ssNo, snNo, nNo), where the "Send" and the "No" represent cases when a cooperative or non-cooperative announcement is made. Similarly, for the committee the memory-one histories become ($\sigma_{\emptyset}$Send, σ_{ss}Send, σ_{sn}Send, σ_{ns}Send, σ_{nn}Send, $\sigma_{\emptyset}$No, σ_{ss}No, σ_{sn}No, σ_{ns}No, σ_{nn}No). We estimate the probability σ after each history and the results are shown in Tables (3) and (4) of Table 2.5.

According to table (1) Table 2.5, while the leniency is low in all treatments, the public is more lenient in treatment CA than other treatments, as the probability to trust after the committee playing "Not Send" is $\sigma_{sn} = 0.41$, which is higher than other treatments. Looking at σ_n, the low probabilities show that once the public decides not to trust, it is difficult for them to trust

[18]The probability of playing "Send" after each memory-one history one the decision-maker's level could be found in Appendix B.3. Overall, the main picture does not change.

Table 2.5: Cooperation rates after memory-one histories

	$\sigma_\emptyset$	σ_{ss}	σ_{sn}	σ_n	ln$L$
I	0.76	0.88	0.24	0.27	-118.65
	(0.12)	(0.05)	(0.08)	(0.15)	
C	0.81	0.95	0.12	0.33	-197.81
	(0.06)	(0.02)	(0.07)	(0.06)	
CST	0.86	0.96	0.22	0.43	-180.30
	(0.06)	(0.03)	(0.07)	(0.12)	
CA	0.87	0.94	0.41	0.51	-195.85
	(0.05)	(0.02)	(0.10)	(0.07)	

(1) public

	$\sigma_\emptyset$	σ_{ss}	σ_{sn}	σ_{ns}	σ_{nn}	ln$L$
I	0.84	0.84	0.65	0.81	0.64	-118.60
	(0.05)	(0.03)	(0.11)	(0.06)	(0.09)	
C	0.69	0.84	0.39	0.72	0.58	-272.099
	(0.07)	(0.02)	(0.07)	(0.07)	(0.04)	
CST	0.77	0.80	0.48	0.77	0.56	-268.99
	(0.05)	(0.03)	(0.08)	(0.06)	(0.08)	
CA	0.78	0.82	0.47	0.77	0.52	-260.12
	(0.05)	(0.04)	(0.10)	(0.06)	(0.11)	

(2) committee

	$\sigma_\emptyset$Send	σ_{ss}Send	σ_{sn}Send	σ_nSend	$\sigma_\emptyset$No	σ_{ss}No	σ_{sn}No	σ_nNo	ln$L$
CA	0.93	0.94	0.46	0.58	0.30	1.00	0.20	0.32	-181.23
	(0.05)	(0.03)	(0.12)	(0.09)	(0.14)	-	(0.18)	(0.10)	

(3) public of treatment CA

	$\sigma_\emptyset$Send	σ_{ss}Send	σ_{sn}Send	σ_{ns}Send	σ_{nn}Send	$\sigma_\emptyset$No	σ_{ss}No	σ_{sn}No	σ_{ns}No	σ_{nn}No	ln$L$
CA	0.86	0.82	0.54	0.76	0.59	0.10	0.60	0.10	0.80	0.40	-242.52
	(0.04)	(0.03)	(0.08)	(0.06)	(0.07)	(0.10)	(0.65)	(0.26)	(0.33)	(0.29)	

(4) committee of treatment CA

Note: Estimated probabilities of playing "Send" after memory-one histories. The four sub-tables from top to bottom are the estimated results for public, committee, and public in treatment CA conditional on announcement and committee choices in treatment CA conditional on announcement. The numbers in parentheses are the standard errors. ln$L$ is the log likelihood of the model.

again. One exception is $\sigma_n = 0.51$ in CA. There are still approximately 50% of the cases in which trust is restored.

For the committees of all treatments in Table (2), if they did not send in the last period, the chance of playing “Send” in the current period is much lower than if “Send” has been played in the previous period. This result is independent of the choices of the public, as the probabilities of sending after history ns and ss are equally high across treatments. This further indicates that the committees are lenient towards the public in the sense that if the public did not send in the previous round, the committee nevertheless sends if they did so in the last round. Interestingly, the individual decision-makers in treatment I send with around 65% probability, even if the committee did not in the previous round. This corresponds to our findings in the previous section, in that there is a last-round effect in treatment I that the rate of sending increases at the starting period.

Table (3) shows that the trust rate of the public is low for all histories if the committee announces not sending. However, once cooperation is established, the public continues to trust. Moreover, the public is more lenient towards the committee if the announcement is “Send”, as the rate of sending σ_{sn}Send in Table (3) is much larger than σ_{sn} of treatment C in Table (3). Table (4) shows that the committee increases the rate of sending if it announces to do so. This effect exists in the initial round and after sn. Comparing $\emptyset$ and σ_{sn} of C in Table (2) with $\emptyset$Send and σ_{sn}Send in Table (4) shows that announcements on the one hand help to establish cooperation in the first round and, on the other hand, helps to restore cooperation.

Result 7: A cooperative announcement makes the public more lenient by making the committee switch back to playing “Send”. It also makes cooperation easier in the initial rounds of every supergame.

To look beyond memory-one and to explain why the play in some treatments are more lenient than others, we next focus on the specific strategies of the public.

2.5.5 Strategies

We estimate strategies for the public rather than for the committee for three reasons. First, strategies specify actions after each possible history. To differentiate among strategies, we need sufficient observations of a sequence of actions following different histories. It is thus difficult to estimate what strategies each decision-maker plays because they interact with a certain public for a maximum of three rounds, which does not provide enough repetitions for estimation. Second, estimating strategies with observations of only three rounds restricts our candidate strategies to no more than three transitions among states. Finally, although committees interact indefinitely with the public, since the decision-makers within the committees change over time, and the committee's choices are the voting results of all decision-makers, a consistent pure strategy at the committee level may not even exist.

Table 2.6: Results of strategy frequency estimation for the public

	I	C	CST	CA
ALLS	-	0.23	-	0.30
		(0.11)		(0.13)
ALLN	0.11	0.06	0.06	-
	(0.11)	(0.06)	(0.05)	
GRIM	0.13	-	-	-
	(0.12)			
TFT	-	-	0.30	-
			(0.13)	
NTFTS	0.11	-	0.06	-
	(0.11)		(0.05)	
FN	-	-	-	0.06
				(0.05)
T2	-	0.39	-	0.42
		(0.13)		(0.14)
T3	0.16	0.22	0.31	0.22
	(0.15)	(0.12)	(0.14)	(0.11)
T4	-	0.09	-	
		(0.09)		
T-MULT3	0.49	-	0.28	-
	(0.19)		(0.11)	
γ	0.10	0.14	0.10	0.12
BIC	189.77	440.85	363.31	393.77
ln$L$	-89.39	-213.20	-174.43	-191.10

Note: This table reports the maximum likelihood shares of pure strategies. The estimation procedure assumes constant strategy use across all supergames. γ is the estimated tremble probability, which avoids the likelihood shares of zero when the subjects deviate from a choice pattern. Strategies are selected based on Bayesian Information Criterion. Strategies that attract zero shares are omitted (-). The standard errors are reported in parentheses. Values may not add up to 1 because of rounding.

Tables B.4.1–B.4.3 in Appendix B.4 summarize the list of candidate strategies. All strategies are modeled using finite automata. The first 18 strategies are pure strategies taken from Engle-Warnick and Slonim (2006). These strategies condition on the choices of the committee. The

Table 2.7: Descriptions of the selected strategies

Acronym	Description
ALLS	Always play “Send”.
ALLN	Always play “Not Send”.
GRIM	Play “Send” as long as the committee returns, otherwise play “Not Send” forever.
TFT	Play “Send” as long as the committee returns, otherwise play “Not Send” for one round and return to “Send”.
NTFTS	Start with “Not Send”, and play “Send” in the next round. Keep playing “Send” as long as the committee returns, otherwise go back to “Not Send”.
FN	Start with “Not Send”, then play “Send” forever.
T2	Play “Send” as long as the committee returns, otherwise play “Not Send” for two rounds and return to “Send”.
T3	Play “Send” as long as the committee returns, otherwise play “Not Send” for three rounds and return to “Send”.
T4	Play “Send” as long as the committee returns, otherwise play “Not Send” for four rounds and return to “Send”.
T-MULT3	Play “Send” as long as the committee returns, otherwise play “Not Send” until the end of round 3 and round 6, then return to “Send”.

transition between states occurs in response to what the committee chose in the previous round.

In addition, we add a pure strategy T-MULT3. Subjects who play this strategy send as long as the committee returns, but switch to “Not Send” if the committee does not send back, then sends as well in round 4 and 7. This is a punishment strategy for treatment I and CST. In round 4 and 7, because the committee is changed, and the public returns to “Send”. The corresponding punishment strategy for treatment C and CA is T2, which is a trigger strategy with two periods of punishment. The public sends as long as the committee returns, otherwise plays “Not Send” for two rounds until the committee has been replaced. Afterwards, they return to “Send”.

The estimation method is adopted from Dal Bó and Fréchette (2011). We use the Maximum Likelihood method to estimate the shares of every strategy in the candidate set for each treatment. A subset of strategies that best describe the behavior of the subjects is selected based on the Bayesian Information Criterion (BIC). The results for the prevalence of the strategies are listed in Table 4.5. Table 2.7 lists each of the selected strategies.

Strategies TFT (tit-for-tat), T2 and T3 are trigger strategies with one, two, and three rounds of punishment. While T3 is popular in all treatments, T2 is played only in collectivist organizations with overlapping generations. Once trust is not returned, the public from treatments

C and CA punishes for two rounds, until the committee has been completely replaced. The one-round punishment strategy TFT is played only in treatment CST.

For treatments without overlapping structures, T-MULT3 is popular in both treatments I and CST, but the share is higher in I than in CST. This implies that the public in these treatments expects that the committee will no longer reciprocate once it keeps the funds. One difference between treatment I and CST is that the public in treatment I plays less forgiving strategies. GRIM is a grim-trigger strategy with the starting action being “Send”. Once trust is not returned, the public who plays GRIM stops trusting in the remaining rounds. In treatment CST, the one-shot tit-for-tat strategy TFT is more popular. It is easier to restore trust under a collectivist organizational structure than under an individualistic organizational structure, but the reason behind is unclear. The forgiving strategy NTFTS is played in both I and CST. This is a tit-for-tat type of strategy but starts with “Not Send”, and punishment is triggered by “Send”.

Treatment CA increases trust by increasing the prevalence of ALLS (always send) and decreasing the prevalence of ALLN (always not send). The public in CA also plays FN (first-round not send), which is a variant of ALLS with the initial action being “Not Send”.

Result 8: Punishment strategies are popular in all treatments, but the length of punishment differs. Collectivist structures make the public more forgiving towards a non-cooperative outcome. Announcements sustain trust by making the public play “always Send” more frequent.

2.6 Conclusion

Trust from the outsiders is essential for the success of many organizations, including central banks. The success of central bank policies depends on the expectations of the public. Therefore, it is crucial to understand which institutional structures foster public trust. This study examines the effect of institutional characteristics on trust and cooperation using an indefinitely repeated binary trust game in the lab.

We focus on two institutional dimensions. The first concerns whether the institution is individualistic or collectivist. Collectivist institutions consist of more than one decision-makers. We further differentiate between two specific structures: collectivist institutions with overlapping terms, and collectivist institutions with synchronized terms. Game-theoretic analysis predicts that it is not possible to establish trust in an individualistic institution and in the collectivist institutions with synchronized terms. In collectivist institutions of overlapping terms, there is always one member who has just joined, one member who joined one period ago, and one member who is about to leave. Trust is possible as long as, for decision-makers who are not in the last term, the future gain from cooperation is higher than the gain of defection. Empirically, institutional structures do not affect the average frequency of trust and cooperation. Trust and cooperation rates are high in an individualist structure, as in a collectivist structure. One possible explanation for the high cooperation rates with an individualist structure is the repeated game effect in which subjects reciprocate for future benefits. However, we find that a substantial share of individualistic decision-makers reciprocate the public's trust until the last term. This may imply that there is a high proportion of reciprocators who cooperate even without future gains.

In the second dimension, we introduce a simple announcement stage to a collectivist institution with overlapping terms. Decision-makers vote on two pre-specific non-binding statements: either always returning the trust or always not returning the trust. We find that an announcement to always exhibit cooperative behavior increases public trust by encouraging more frequent use of cooperative strategies. The organization follows through with the announcement. The effect of announcements is short-lived in terms of tenure, but long-lasting over time. Decision-makers in the last term do not follow a cooperative announcement as much as their counterparts in earlier terms. However, the rate of sending back the transfer on the organization's level after a cooperative announcement remain high and stable over time, because the composition of the organization is stable over time.

It is important to note that decision-makers are limited to choosing from two pre-determined messages. One of the statements, that the organization will always not return the trust, is not popular and does not make sense to many decision-makers. If decision-makers were able to create their own messages, these could have even stronger consequences for behavior. Therefore, our findings regarding the effect of the announcement represent a conservative estimate or a lower bound.

Future research could examine more deeply the behavioral factors underlying the high cooperation rates of the individualistic treatment. It is possible that the diffusion of responsibilities among multiple decision-makers makes the collectivist organization less trust-worthy, and hence leads the public to have a similar level of trust in the individualistic and collectivist organizations. However, our experimental design does not allow us to measure directly the effect of perceived diffusion of responsibility on the trust of the public. Future research could explore more deeply whether and how trust is influenced by the diffusion of responsibilities. Last but not least, it would be interesting for the future research to focus on the effect of announcements on trust by allowing the decision-makers to formulate their own messages.

Chapter 3

Sustaining Cooperation With Correlated Information

With Fabian Dvorak and Sebastian Fehrler

Abstract

In indefinitely repeated games in which the players cannot observe the true actions of the partners but receive noisy information regarding their partners' actions, sustaining cooperation is difficult. Theoretical literature has found that cooperation can be sustained when noisy information is correlated. However, it is difficult to identify correlated information. In this study, we implement a correlated information structure using laboratory experiments. Our experimental results show that a substantial fraction of subjects use a strategy that takes into account of the correlated information. However, correlated information does not result in more cooperative actions, as behavior is more lenient when information is independent.

3.1 Introduction

A large literature in game theory investigates cooperation in indefinitely repeated games. In some cases, players monitor the actions of the opponents perfectly. In many situations, however, opponents' actions are monitored only imperfectly. For example, in teamwork, players repeatedly put effort into a joint project. Their partners do not observe their actual effort, but instead the output which noisily represents the effort. In the oligopoly model of collusion (Green and Porter, 1984), firms observe the noisy market prices. In both examples, shocks or other factors may confound the observed information, such that instead of directly observing the actual choice of the opponent, players only observe noisy signals that represent the opponents' actions.

Understanding how cooperation can be sustained in indefinitely repeated games under imperfect monitoring has grown to an important strand of literature, because sustaining cooperation is difficult when the monitoring structure is imperfect. Awaya and Krishna (2016) point out that one way to sustain cooperation is through correlated signals. While the literature commonly assumes that noisy signals solely depend on players' own chosen actions, Awaya and Krishna (2016) assume that the noisy signals depend on all actions. With their assumption, the revealed signals are correlated in a way that they are similar when players' actions are similar, and become dissimilar, as players' actions diverge. One example is the Duopoly. The two firms have private information regarding sales, but their sales are correlated as they are sensitive to the prices on the market. The fundamental insight in Awaya and Krishna (2016) is that exploiting correlation in information can help to improve monitoring and sustain cooperation. In this study, we adopt and adapt their assumption of information correlation and explore the empirical implications.

Intuitively, consider a situation when signals are perfectly correlated, i.e. they are the same when actions are the same, and are independently drawn when actions diverge. When signals are only privately observable, communication is necessary to detect correlation. When they are publicly observable, players can detect deviation from the signals to achieve cooperation. Our focus is on the second scenario. In such cases, individual strategies may depend on the correlation of the signals. Cooperation can be enforced by a grim-trigger strategy which conditions on signal correlation. For instance, if both players start with cooperation, both will receive the same signal, and will continue cooperating. Once the signals become different, they infer from the dissimilarity of the signals that different actions have been chosen in this period. The cooperating player detects deviation and defects as well for future periods. As this strategy is similar to grim-trigger strategy except that it takes the correlation of information into account, we refer to this strategy as the correlated grim-trigger strategy (CGRIM). This paper considers other correlation-variational strategies such the correlated tit-for-tat (CTFT), the correlated win-stay-lose-shift and the trigger strategy with two periods of punishment (CT2). We show theoretically that in our case coordination on CGRIM would lead to full cooperation while zero cooperation is possible without correlation.

Our main finding is that subjects use the correlation-variational strategies when signals are perfectly correlated. They use in particular CGRIM and CTFT. However, we do not find a cooperation promoting effect of correlation as 1) CGRIM triggers many punishments, and 2) subjects' play is very lenient in the absence of correlation.

We implement an indefinitely repeated Prisoners' Dilemma of imperfect public monitoring, where payoffs depend on the own action and the signal regarding the action of the other player. Signals are publicly observable and noisy, i.e. they falsely represent the real action of the opponent with an exogenously given and fixed probability. We distinguish between two settings. In one experimental treatment we implement a correlation structure which can be exploited to support full cooperation as subgame perfect equilibrium. Signals are systematically and perfectly correlated, i.e. the two public signals are the same, if both players choose the same action. If their actions differ, signals are independently drawn. We compare subjects' decisions in this treatment to a control treatment with independent signals. The control treatment has a similar design as Fudenberg et al. (2012); Aoyagi et al. (2019); Dvorak and Fehrler (2018). In the first treatment, signals have an improved quality, as they reduce the uncertainty towards the opponent's action. The correlation structure of signals should affect players willingness to cooperate, because players can infer actions from signals when signals are perfectly correlated, and as a consequence, detect deviation of the opponent. We choose parameters such that the cooperative equilibrium can be supported with the correlation treatment but not without correlation, and we expect to observe a higher cooperation rate with signal correlation. However, it is not trivial whether people are capable of detecting correlation. Theoretical and empirical evidence shows that people neglect information correlation when forming beliefs (Enke and Zimmermann, 2017; Levy and Razin, 2015). This question cannot be easily answered from observational data. Because relevant information about the signals and their correlation structure is usually unknown and cannot be exogenously manipulated. Moreover, due to the existence of multiple equilibira, it is not surprising when the non-cooperating strategies are played. The treatments have been pre-registered on the AEA RCT Registry. Our pre-analysis plan is in the Appendix C.6. [1]

Under imperfect monitoring, bad signals occur with positive probabilities even when players cooperate. Players may deviate from cooperation by mistake if they deviate upon bad signals. Efficient equilibria require lenient and/or forgiving strategies. This means that players do not revert to defection immediately after receiving a defective signal (lenient) and are willing to return back to cooperation after punishing defection (forgiving). In the signal correlation treatment, we expect subjects to adjust their strategies accordingly to be conditional on joint signals. However, such adjustment is likely to shift the strategies towards lower leniency and forgiveness, because when it is possible to infer the real action of the opponent from the correlated signals, an efficient outcome can be achieved without coordination on lenient and forgiving strategies. We will explore in more detail whether leniency and forgiveness differ between treatments.

We implement pre-play communication before each interaction in a form of free chat. Communication facilitates cooperation (e.g. Cooper et al. 1992; Rabin 1994; Ellingsen and Östling 2010). Pre-play communication, in particular, helps players to coordinate on cooperative equilibrium, and raises cooperation rates by reducing strategic uncertainty (Kartal and Müller, 2018). In a previous study of a noisy, indefinitely repeated Prisoners' Dilemmas with uncorrelated signals, Dvorak and Fehrler (2018) observe high cooperation rates in the first round with pre-play communication followed by a steady and substantial decline in the subsequent rounds. Communication is not a treatment variable in our design, but it gives us a higher chance of observing cooperative strategies, and allows for the comparison of the decline of cooperation rates between treatments. We further classify the communication content and analyze its impact on the evolution of strategy choices.

To better understand subjects' choices of strategies, we build on the strategy frequency estimation method (SFEM) introduced by Dal Bó and Fréchette (2011) that investigates how strategies differ between treatments. The SFEM is frequently used to obtain maximum-likelihood estimates on the shares of each strategy from a candidate strategy set with experimental data.

[1] In our contingent pre-analysis plan, we intended to implement up to five treatments. But we had to stop after the second treatment because we observed no significant treatment difference in the average frequency of cooperation. Detail can be found in the Appendix C.6.

We estimate the strategy shares using the pure strategies in Dal Bó and Fréchette (2011) and their correlation-variants CGRIM, CTFT, CWSLS and CT2 as candidate set. For robustness check, we further include the strategies from Fudenberg et al. (2012). We find that the non-lenient and non-forgiving CGRIM is the most popular strategy in the signal correlation treatment. In addition, subjects play CTFT. The strategies in the treatment without signal correlation does not condition on the correlation of the signals, and ALLC is the mostly adopted strategy, with a frequency of 73 percent. By contrast, ALLC attracts only 24 percent of the likelihood shares in the treatment with correlation. The results show that 1) subjects do not ignore correlation when signals are correlated, and that 2) subjects in the treatment without correlation are more cooperative and the play is more lenient. In terms of forgiveness, cooperation will be restored from the punishment phase when subjects play CTFT, whereas CGRIM prescribes the strongest punishment. Our estimation shows that on average, the willingness to return back to cooperation is equally low in both treatments, i.e. 3) there is no difference in the level of forgiveness between treatments. Our further analysis shows that improving signal quality with correlation does not significantly affect cooperation rates. The average frequency of cooperation is 78 percent in the signal correlation treatment, and 79 percent in the other. The cooperation rates start high in both treatments and decline over rounds. The correlation treatment has a slightly steeper decline of 10 percent than the treatment without correlation. These observations can be explained by subjects' choice of strategies. Subjects are much more lenient when signals are not correlated; since many subjects follow strategies other than CGRIM, there is a high probability that many punishments will be triggered.

Our study is motivated by Awaya and Krishna (2016; 2019) with one major difference in the monitoring structure. While their focus is on private monitoring, our design is of public monitoring.[2] They show that communication prior to every round can sustain cooperation based on pure strategies. A necessary condition is that the private signals are more strongly correlated if players choose the same action. By publicly reporting private signals, players can draw inferences about past actions based on these reports. However, under public monitoring, as signals are publicly observable, no useful information will be exchanged through reporting. When players communicate truthfully their signal under private monitoring, a public history is

[2]Private monitoring, as in Sekiguchi (1997); Compte and Postlewaite (2015), means that noisy signals are privately observable. The secret price cutting of oligopolies described by Stigler (1964) is a classic example.

created. Our implementation represents the optimal case with truthful communication under private monitoring, and allows for simpler and stable equilibria.[3]

Our central focus is thus not on the reporting strategies players devise when they know how to exploit information contained in correlation, but rather on whether they make use of such correlated information in choice making given that everyone reports truthfully. Our design does not answer directly the research question of Awaya and Krishna (2016), i.e. what effect does the exchange of private information have on cooperation given signal correlation. We instead test the underpinning of their theory by addressing the research question to what extent subjects can exploit information contained in correlation.

The rest of the paper is organized as follows. In the next section, we introduce the stage-game, and derive theoretical predictions of both treatments. In section 3, we describe the experimental design and clarify the reasons for our choice of stage-game parameters. Section 4 summarizes research questions and methodology. We present the empirical results in section 5. The key findings and methodology are summarized in section 6.

3.2 Repeated Prisoners' Dilemma with Imperfect Public Monitoring

Consider an indefinitely repeated Prisoners' Dilemma with two players who repeatedly play a 2×2 stage game. The game terminates with a constant exogenously given probability of $0 < 1 - \delta < 1$ after every round. In each round, the two players $i = \{1, 2\}$ choose from an action set $A_i = \{C, D\}$ and their action is denoted by a_i.

There is ex-post uncretainty regarding the actual choice of an opponent. Under public monitoring (Green and Porter, 1984), instead of observing the actual choice, each player's action is translated into a noisy signal and both signals are publicly announced to players. More technically, let $\Omega_i \in \{c, d\}$ be a set of signals of player i's action, where signal c represents action

[3] Under private monitoring, the mixed strategy equilibria that support cooperation are difficult to implement and evolutionary unstable (Compte and Postlewaite, 2015; Heller, 2017). However, when players have a chance to communicate their private signals, truthful communication equilibria exist by satisfying certain revelation constraints. If signals are correlated, truthful communication equilibria with strict revelation constraints can be constructed (Kandori and Matsushima, 1998). The cooperative truthful communication equilibria are shown to be not only weakly stable, but also survives the evolutionary stability criterion (Heller, 2017).

C, and d represents D. We denote the realized signal from this set by ω_i. For each action profile $a = (a_i, a_{-i})$, a conditional probability distribution $\pi(\omega|a)$ is assigned over signals. In our experiment, the two treatments differ with respect to π, and therefore in the quality of signals.

Imperfect public monitoring with uncorrelated signals (NoCor) For the treatment without correlation, signals are conditionally independent of $a = (a_1, a_2)$. That is, signals are drawn independently for each of the chosen actions and signal ω_i depends only on a_i but not on a_{-i}. Signals are noisy and indicate the opposite of the chosen action with probability $\epsilon = 0.2$, that is: when the player chooses C (D), the signal indicates C (D) with probability $1 - \epsilon = 0.8$, and indicates D (C) with probability 0.2. The conditional probability distribution of ω_i is thus $\pi(\omega_i = c|a_{-i} = C) = \pi(\omega_i = d|a_{-i} = D) = 1 - \epsilon$ and $\pi(\omega_i = c|a_{-i} = D) = \pi(\omega_i = d|a_{-i} = C) = \epsilon$.

Imperfect public monitoring with correlated signals (Cor) In this setup, both players receive identical signals if they choose the same action. Yet the signals still indicate the opposite of the chosen action with probability 0.2. The conditional probability distribution of ω_i is given by $\pi(\omega_i = c|a_i = a_{-i} = C) = \pi(\omega_i = d|a_i = a_{-i} = D) = 1 - \epsilon$ and $\pi(\omega_i = c|a_i = a_{-i} = D) = \pi(\omega_i = d|a_i = a_{-i} = C) = \epsilon$ when both players choose the same action. If their actions differ, signals are drawn independently, as in NoCor. Therefore, if the two signals differ, it is for certain that different actions are chosen. If the signals are the same, players do not know for sure the opponent's action.

Player i's stage game payoff g_i is defined by i's action a_i and the noisy signal ω_{-i} reflecting the other player's action, i.e. $g_i : A_i \times \Omega_{-i} \to \mathbb{R}$. Hence player i cannot infer the action of the other player from her payoff. The expected stage game payoff of player i is given by $u_i : A \to \mathbb{R}$, $u_i(a) = \sum_{\omega_{-i} \in \Omega_{-i}} g_i(a_i, \omega_{-i}) \pi(\omega_{-i}|a)$.

The expected stage-game payoff profile (u_i, u_{-i}) which we consider has the form of a Prisoners' Dilemma:

	C	D
C	$1, 1$	$-l, 1+g$
D	$1+g, -l$	$0, 0$

Expected payoffs in the above table are normalized. Parameters g and l are both positive and $g < 1 + l$.

Denote by $\Omega^t = \{\Omega_i, \Omega_{-i}\}^t$ the set of public histories up to round t . A public strategy for player i is a mapping $\sigma_i : \bigcup_{t\geq 0} \Omega^t \rightarrow \Delta A_i$. A strategy profile $\sigma = (\sigma_i, \sigma_{-i})$ is a public perfect equilibrium (PPE) if σ_i is a public strategy and for any history up to round t, σ is a Nash equilibrium in all rounds following t, in other words PPE is a subgame perfect Nash Equilibrium (SPE) that depends only on public signals.

The equilibrium solution for NoCor Mutual cooperation can be sustained by non-lenient and non-forgiving grim-trigger strategies. Players start with C but deviate to D for all subsequent rounds if the private history $(a_i, \omega_i, \omega_{-i})^t \neq (C, c, c)^t$. We refer to mutual cooperation as the reward state, and to mutual defection as the punishment state. This strategy is a PPE if the long-run incentive of cooperation $\frac{1}{1-\delta(1-\epsilon)^2}$ is as least as big as the gain from defection $\frac{1+g}{1-\delta\epsilon(1-\epsilon)}$, i.e. δ should be sufficiently large

$$\delta \geq \delta^{PPE}_{NoCor} = \frac{g}{(1+g)(1-\epsilon)^2 + \epsilon^2 - \epsilon} \tag{3.1}$$

The equilibrium solution for Cor Mutual cooperation can still be enforced by grim-trigger strategies. The class of cooperative SPE strategies defined in the above paragraph still exists. There additionally exist correlated grim-trigger strategies which enforce cooperation. Specifically, we define a correlated grim-trigger strategy as follows: both players start with C and will continue choosing C as long as the two signals match, i.e. $\omega_i = \omega_{-i}$. Otherwise revert to D in all subsequent rounds. Cooperation is optimal if incentive of cooperation is at least as big as that of defection. The threshold δ^{PPE} is given by

$$\delta \geq \delta^{PPE}_{Cor} = \frac{g}{1 - 2\epsilon + 2\epsilon^2 + g} \tag{3.2}$$

See Appendix C.1 for proof.

It is easier to sustain cooperation under Cor than NoCor in terms of grim-trigger strategies. For $\epsilon \in (0, 1)$ and g fixed, whenever (3.1) holds, condition (3.2) holds as well. The correlation structure of signals changes the equilibrium condition in two aspects. First, it increases the value in the continuation path of the reward state. Even when both players cooperate, bad signals occur with positive probability in NoCor which triggers the punishment state. Under NoCor, the continuation value on the cooperative path is the expected utility of reward state and punishment state. With correlation, however, signals always match if both cooperate, which means punishment state will never be reached, and therefore the continuation value is larger.

Second, note that Cor makes unilateral defection more attractive as the probability of staying in reward state after unilateral defection is higher than that in NoCor. When signals differ, a cooperating player can know for sure that the opponent has defected. However, when signals match, it is impossible to make such inference. Even when players choose different actions, signals still match with a probability of $2\epsilon(1-\epsilon)$, i.e. the probability of remaining in the reward state after unilateral defection is the probability of the signal profile being either $(\omega_i, \omega_{-i}) = (c, c)$ or (d, d), that is, $2\epsilon(1 - \epsilon)$. The NoCor treatment renders this probability to be $\epsilon(1 - \epsilon)$, i.e. cooperation in the next round is only possible when (c, c) is observed. Despite that NoCor triggers punishment with a higher probability when one player defects unilaterally, it does not compensate for the reduced expected utility in the reward state due to punishments on the equilibrium path.

3.3 Experimental Design

We implement a noisy Prisoners' Dilemma game with public monitoring in a laboratory experiment. The experiment has two between-subject treatments. The two treatments vary in the conditional distribution of signals:

NoCor Signals are public and independent.

Cor Signals are public and perfectly correlated if both actions are the same, and independent otherwise.

In every round, two players choose their actions $a_i \in \{C, D\}$ simultaneously. Payoffs depend on the player's own action a_i and the received signal about the other player's action $\omega_{-i} \in \{c, d\}$. Signals are noisy and indicate the wrong action with probability $\epsilon = 0.2$, which do not vary between treatments.

Continuation probability δ of the repeated game is 0.8. Stage-game payoff matrix is given in the upper panel of Figure 3.1. The payoffs are in experimental currency units. The lower panel of Figure reffig:stage-game-parameters shows the expected stage-game payoffs. The normalized expected stage-game parameters are $g = l = 0.8$.

Figure 3.1: Stage-Game Parameters and Predictors of Cooperation

	c	d
C	32	2
D	40	10

	C	D
C	$26, 26$	$8, 34$
D	$34, 8$	$16, 16$

Notes: Payoffs are in experimental currency units with an exchange rate of 50 ECU = 1 EUR. Both matrices are shown to subjects during their decision-making.

Under such parameterization, the existence thresholds of PPE are $\delta^{PPE}_{NoCor} = 0.81$ and $\delta^{PPE}_{Cor} = 0.54$. These parameters make sure that condition (3.2) holds but (3.1) does not, i.e. there exists an equilibrium in which both players play correlated grim-trigger strategy when signals are correlated, and no cooperative PPE exists if there is no correlation.

In all treatments, subjects engage in a pre-play communication-stage before the first round of every supergame. A supergame is an infinitely repeated interaction. Communication is possible via a chat-box interface for 120 seconds.

In every session of both treatments, subjects are randomly assigned to 3 matching groups with 8 participants in each group. Subjects only interact with the members in the same matching group throughout the entire session. Thus, one matching group can be treated as one independent observation. Subjects play 7 supergames of pre-determined lengths. At the beginning of every supergame, subjects are matched with a new partner from their matching group using perfect stranger matching such that they do not play with the same partner from the matching group for a second time. To keep the length of supergames constant across treatments, we generate 3 sequences of random numbers beforehand, and use them to determine the length L_i of each supergame.[4] To increase the number of observations per supergame, we adapt the block-random-termination method (Fréchette and Yuksel, 2017). Subjects play a block of five rounds at the beginning of every supergame. If the true length L_i is smaller or equal than 5, the supergame ends at the end of round 5 and only the first L_i rounds are payoff relevant. If L_i is larger than 5, the supergame continues until round L_i has been reached and all rounds are payoff relevant. Before the end of round 5, subjects are not informed about whether the supergame ends or not. The modified block-random-termination method allows us to collect data of at least five rounds in every supergame.

At the end of the first four rounds within each supergame, subjects receive feedback on $\{a_i, \omega_i, \omega_{-i}\}$ plus the stage profit g_i. Starting from round 5, in addition to $\{a_i, \omega_i, \omega_{-i}\}$ and g_i, they are informed of the pre-generated random number of that round. The supergame does not end until that number is smaller or equal than $1-\delta$, which is 0.2 under our parameterization.

Before the game starts, subjects are informed about each detail of the experimental implementation and required to answer control questions before the game starts to make sure they understand the experimental procedure. At the end of the experiment, subjects answer a short survey that elicits basic socio-economic characteristics, such as age and gender.

We conducted six sessions in February 2020 at Lakelab of University Konstanz, with three sessions for Cor and NoCor respectively.[5] A total of 144 subjects participated in our experiment. Subjects were students and employees of the University of Konstanz, Germany. After data collection, we found out that one subject participated twice in two NoCor sessions. We exclude the matching group of this subject's second participation from data analysis. Each session lasted 2 hours, and the average earning was 18.67 Euros.

[4] We use Stata to generate 3 sequences of uniformly distributed random numbers between 0 and 1 with seeds 3, 4, and 5 (Seeds 1 and 2 have been used in: Dvorak and Fehrler, 2018). Denote the 3 sequences as $\{r_n\}_i = \{r_1, r_2, ..., r_x\}_i$, where $i = 3, 4, 5$ indicates the seed underlying the sequence and $n \in \mathbb{N}$. The first supergame has x_1 rounds if $r_{x_1} \leq 0.2$ and for all $n < x_1$, $r_n > 0.2$. The second supergame has $x_2 - x_1$ rounds if $r_{x_2} \leq 0.2$ and for all $x_1 < n < x_2$, $r_n > 0.2$. And so forth. The resulting (lengths of the) sequences are SQ1 $(2, 8, 1, 5, 7, 1, 7)$, SQ2 $(4, 2, 2, 21, 4, 3, 5)$ and SQ3 $(2, 3, 1, 1, 4, 6, 6)$.

[5] The experiment was programmed in z-Tree (Fischbacher, 2007) and subjects were recruited via hroot (Bock et al., 2014).

3.4 Research Questions

We address three research questions on the treatment difference in average strategy choice, shares of strategies, cooperation rates, and the effect of communication on cooperation.

Question 1: *How do strategy choices differ with different informational correlation structure?*

We will take an in-depth examination on the cooperation rates after memory-one histories and the individual heterogenous strategy choices to shed light on the treatment difference in leniency and forgiveness. We are particularly interested in assessing shares of CGRIM strategy.

In our setup, a memory-one history consists of own action and the two public signals $\{a_i,\omega_{-i},\omega_i\}$. Therefore, we have a vector of nine possible memory-one states ($\emptyset$, ccc,ccd,cdc,cdd,dcc,dcd,ddc,ddd). The first element $\emptyset$ is the initial round with no history of play. The rest of the elements are nonempty histories representing $\{a_i,\omega_{-i},\omega_i\}$. For instance, cdc describes a state when a focal player chooses cooperation, receives signal d and sends out signal c to the other player. We look at the cooperation probabilities after each of the possible histories and we represent the cooperation probabilities by a vector ($\sigma_\emptyset$, σ_{ccc},σ_{ccd},σ_{cdc},σ_{cdd},σ_{dcc},σ_{dcd},σ_{ddc},σ_{ddd}). We focus on nine states to allow behavior to be conditioned on the combination of action and public signals. This allows us specify when defection should be triggered when estimating the shares of correlation-variational strategies. Unjustifiable defection is estimated by $1-\sigma_{ccc}$, $1-\sigma_{ccd}$ without signal correlation, and by $1-\sigma_{ccc}$, $1-\sigma_{cdd}$ with signal correlation. σ_{cdc}, σ_{cdd} are estimations for leniency when signals are uncorrelated. When signals are correlated, estimations for leniency are σ_{ccd} and σ_{cdc}. For both uncorrelated and correlated signals, σ_{dcc}, σ_{dcd}, σ_{ddc} and σ_{ddd} are estimations for forgiveness.

To analyze strategy choices, we use the R package stratEst developed by Dvorak (2018), which adapts and extends the SFEM of Dal Bó and Fréchette (2011) using EM algorithm (Dempster et al., 1977). We restrict our attention to the first five rounds of the last 3 supergames when subjects have gained experience of play. In the main text, we report models which include the six pure strategies studied by Dal Bó and Fréchette (2011) and their corresponding correlation-variants as our candidate strategies. We adapt these pure strategies such that they condition on the own action and the public signals. Descriptions of the candidate strategies are summarized in Table 3.1.[6] For robustness check, we extend the candidate set to include the 20 strategies in Fudenberg et al. (2012). The comprehensive description and the automata of all strategies are in the tables of Appendix C.2. We report the strategy estimation results with all 20 pure strategies and the four correlation-variants in Table C.3.1 of Appendix C.3.

Table 3.1: Overview of the Candidate Strategies

Strategy	Acronym	Description
Always Defect	ALLD	Always play D.
Always Cooperate	ALLC	Always play C.
Grim	GRIM	Play C until the signal representing the partner's action is d, then play D forever.
Tit-for-Tat	TFT	Play C unless the signal representing the partner's action is d in the last round.
Win-Stay-Lose-Shift	WSLS	Play C if own choice is the same as the signal representing the patner's choice in the last round, otherwise play D.
T2	T2	Play C until either signal is d, then play D twice and return to C (regardless of all actions and signals during the punishment rounds).
Correlated Grim	CGRIM	Play C until the public signals do not match, then play D forever.
Correlated Tit-for-Tat	CTFT	Play C unless the public signals do not match in the last round.
Correlated Win-Stay-Lose-Shift	CWSLS	Play C if the public signals match, otherwise play D.
Correlated T2	CT2	Play C until the signals do not match, then play D twice and return to C (regardless of all actions and signals during the punishment rounds).

Question 2: *Is the level of cooperation higher in Cor than in NoCor?*

According to the theoretical prediction outlined in Section 2, there should be more cooperation in the Cor treatment. However, if subjects ignore correlation, their inability to exploit correlation will predict the opposite. We formulate our main hypothesis in line with the theory:

[6] ALLC and ALLD do not condition on past histories, but can be represented by a nine-state vector, i.e. they are defined respectively by $(1, 1, 1, 1, 1, 1, 1, 1, 1)$ and $(0, 0, 0, 0, 0, 0, 0, 0, 0)$.

H1: *The average cooperation rate will be higher in Cor than in NoCor.*

H1 follows directly from condition (1) and (2). We test this hypothesis by comparing the average cooperation rates in the first five rounds of the last three supergames between treatments. We further turn to the decline of cooperation rates to compare the stability of cooperation.

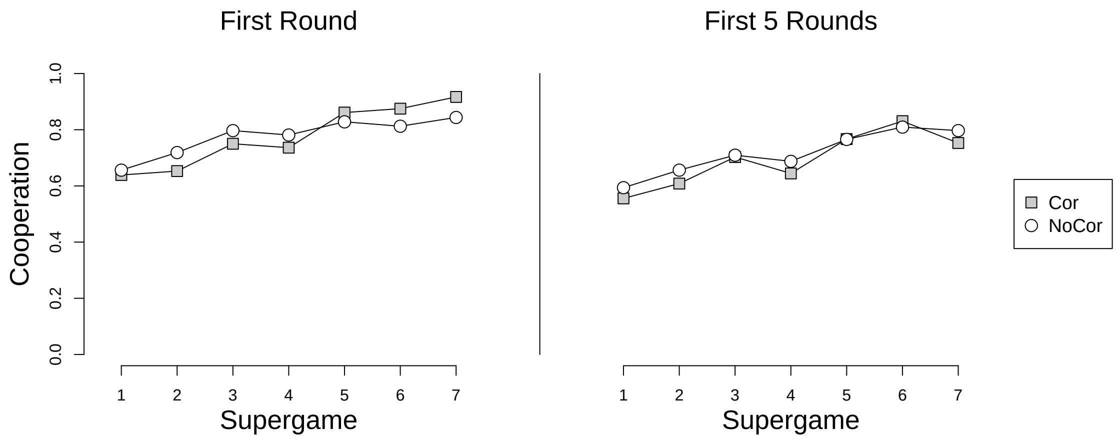
Question 3: *What do subjects talk about? How does communication affect cooperation in general? And how does communication affect strategy choices?*

To approach the problem, we have subjects' communication content classified by two research assistants into 39 sub-categories, which we then group into four main categories. Tables in Appendix C.4 document in detail these categories. We expect those who coordinate in the Cor treatment communicate either explicitly or implicitly about punishing their opponent in case the public signals differ. Our major focus is on the main categorical level, which we take as explanatory variables and analyze through logistic regressions whether subjects who talk about certain topics are more likely to play cooperatively. We take a further look at the individual strategy choices using the frequency of communication in the main categories on subject level as covariates. The method is developed by Dvorak and Fehrler (2018) in the spirit of latent-class regression models (Dayton and Macready, 1988; Bandeen-Roche et al., 1997).

3.5 Experimental Results

It takes a few supergames for subjects to learn and stabilize their behavior. Using the unstabilized choices from the first few supergames violates the basic assumption for strategy estimation that the distribution of strategies should be stable across supergames. We therefore proposed in our pre-analysis plan to focus on the last three supergames for cooperation rates and strategy frequency analysis. Figure 3.2 depicts the average cooperation frequency over all seven supergames. The average cooperation rates stabilizes as subjects gain more experience. In addition, since each supergame consists of at least five rounds, our main focus is on the first five rounds in the supergames such that the transition from the block to random termination does not confound the result.

Figure 3.2: Evolution of Cooperation over Supergames



Notes: The lines depict the first round and average frequency of cooperation over seven supergames for both treatments.

3.5.1 Strategy Choices

We estimate the probability of cooperation after different memory-one histories. For both treatments, cooperation rates condition on $\{a_i, \omega_{-i}, \omega_i\}$, which has nine possible memory-one histories. Table 3.2 reports the cooperation rates following each history. $\sigma_\emptyset$ indicates the probability of cooperating in the first round. It is high in both treatments, meaning that subjects are generally cooperative in the initial round. The unjustified defection ($1 - \sigma_{ccc}$, $1 - \sigma_{cdd}$ in Cor and $1 - \sigma_{ccc}$, $1 - \sigma_{ccd}$ in NoCor) rates are low in both treatments. The largest treatment difference is observed within the *cdc* state. Subjects in NoCor are substantially more cooperative than in Cor. Since σ_{cdc} is an estimation of leniency in both treatments, a higher level indicates that subjects are less likely to defect when a bad signal occurs. State *ccd* and *cdd* are also estimations of leniency in Cor and NoCor respectively. Treatment differences between these states are small. We do not observe much difference in states *dcc* and *dcd*. The cooperation rates in these states are relatively low in both treatments. In terms of forgiveness (*dcc*, *dcd*. *ddc* and *ddd*), the willingness to return to cooperation after mutual defection is comparably low in both Cor and NoCor.

The probabilities of ending in these states are unequal for all strategies. We therefore cannot interpret the cooperation probabilities after memory-one histories as strategy choices and explain whether subjects play a certain strategy or not. For instance, σ_{ccd} and σ_{cdc} are very different in the correlation treatment. Subjects are willing to cooperate when they receive a cooperative signal from their partners but tend to defect when the signal regarding the partner's choice is defective. If subjects play CGRIM, σ_{ccd} and σ_{cdc} should equalize. But we still cannot conclude that subjects in Cor do not play CGRIM due to the above reason. To find out what strategies they play, our next step is to investigate the heterogeneity of strategy choices.

Table 3.2: Cooperation Rates After Memory-One Histories

	$\sigma_{\emptyset}$	σ_{ccc}	σ_{ccd}	σ_{cdc}	σ_{cdd}	σ_{dcc}	σ_{dcd}	σ_{ddc}	σ_{ddd}	ln$L$
Cor	0.88	0.95	0.83	0.33	0.71	0.41	0.39	0.17	0.35	-409.28
	(0.03)	(0.01)	(0.22)	(0.06)	(0.04)	(0.08)	(0.09)	(0.21)	(0.07)	
NoCor	0.83	0.91	0.95	0.79	0.86	0.22	0.46	0.11	0.32	-386.73
	(0.04)	(0.02)	(0.03)	(0.04)	(0.07)	(0.13)	(0.10)	(0.09)	(0.09)	

Notes: This table summarizes the response probabilities of cooperation following the 9 possible histories: $\emptyset$, *ccc*, *ccd*, *cdc*, *cdd*, *dcc*, *dcd*, *ddc*, *ddd*. Bootstrapped standard errors are from 10000 iterations and are presented in parentheses. The log likelihood of the model is summarized in the last column.

We are interested in the specific strategies subjects employ, which contribute to our observation in Table 3.2 and explain why subjects in NoCor are substantially more lenient on average following *cdc*. Similarly, all candidate strategies condition on $\{a_i, \omega_{-i}, \omega_i\}$. We follow the strategy estimation procedure of Dal Bó and Fréchette (2011). The method estimates a mixture model on candidate strategy set. We select from this given set a subset of strategies that best fits subjects' choices based on Bayesian Information Criterion (BIC). Table 3.3 shows the estimated maximum-likelihood shares of the selected strategies for both treatments. We test whether the strategy shares differ between treatments with a likelihood-ratio test by assuming that both treatments have identical shares for all strategies in the nested model. The $p-$value is < 0.001, which suggests that the strategy shares differ between treatments. In Table 3.3, ALLC attracts a substantial share in NoCor, whereas the non-lenient and non-forgiving CGRIM gets the highest share in Cor. This estimation result explains and confirms the low leniency of Cor following state *cdc*. The behavior of a substantial share of subjects in Cor are consistent with CGRIM that they defect immediately when signals mismatch and never return to cooperation.

To check the robustness of the SFEM result, we expand the candidate strategy set to include another 14 pure strategies. We in addition involve CGRIM, CTFT, CWSLS and CT2 into the candidate set to as the correlation-variational strategy. Table C.3.1 in Appendix C.3 presents

Table 3.3: SFEM with Strategy Selection

	ALLD	ALLC	TFT	WSLS	CGRIM	CTFT	γ	BIC	ln$L$
Cor	0.06	0.26	0.06	0.07	0.40	0.16	0.10	856.59	-415.47
	(0.03)	(0.08)	(0.04)	(0.05)	(0.10)	(0.08)			
NoCor	0.12	0.76	0.07	-	0.05	-	0.10	737.02	-360.19
	(0.04)	(0.06)	(0.04)		(0.03)				

Notes: The table reports the maximum-likelihood shares of the strategies of Dal Bó and Fréchette (2011) and their correlation-variants with data from the first 5 rounds of the last 3 supergames. All strategies condition on action-public signal profile $\{a_i, \omega_{-i}, \omega_i\}$. The estimation procedure assumes constant strategy use. γ is the estimated tremble probability, which avoids likelihood shares of zero when subjects deviate from a choice pattern. Strategies are selected based on Bayesian Information Criterion. Strategies attracting zero shares are omitted (-). Standard errors are reported in parentheses. Values may not add up to one because of rounding.

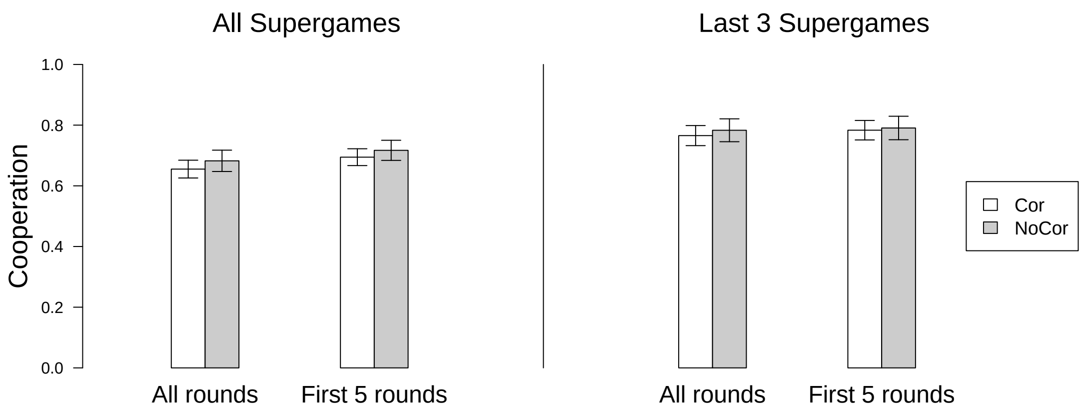
the results. In general, ALLC remains the most popular strategy in NoCor, while in Cor, although CGRIM is still one of the selected strategies, it does not have the highest shares.

Result 1: Subjects are less lenient when signals are correlated, especially when they are in state *cdc*. The cooperation probability of following this state is much lower in Cor. The strategy frequency estimation confirms this observation. Subjects do not ignore signal correlation, but are very cooperative in the NoCor treatment. They play the cooperative and lenient ALLC in NoCor, and play the non-lenient CGRIM in Cor. The difference in strategy choices between treatments is likely to influence the average frequency and stability of cooperation rates, which leads us to our second research question.

3.5.2 Cooperation

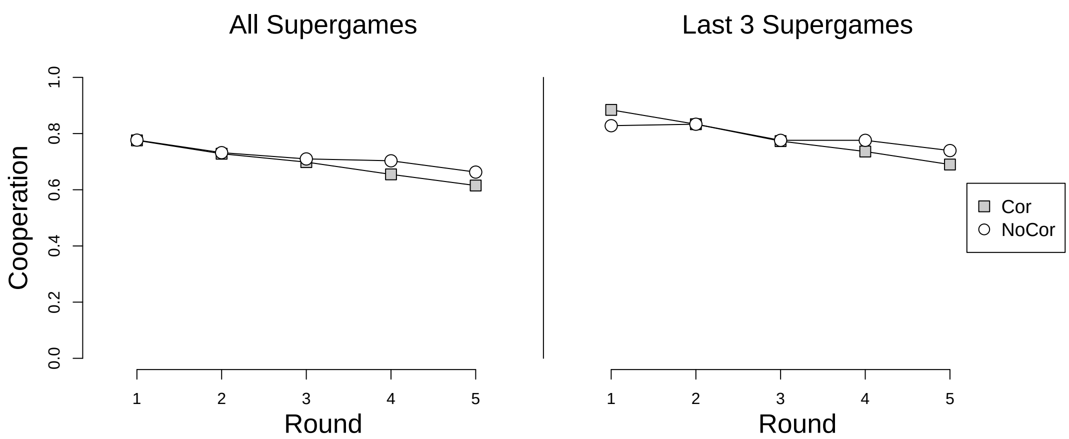
Figure 3.3 shows the average cooperation rates in both treatments. The bars indicate the mean of the cooperation rates, and the whiskers are the two-way clustered standard errors of the mean (Cameron et al., 2011). We estimate the two-way clustered standard errors on participant-match level and calculate the p-value accordingly. The bars indicate that the average cooperation rate does not differ between treatments. The two-sided p-value is 0.44, with an effect size of 0.01 (the average cooperation rate is 0.78 in Cor, 0.79 in NoCor).

Figure 3.3: Average Cooperation Rate Between Treatments



Notes: Bars show the average frequency of cooperation. Whiskers are two-way clustered standard errors for mean cooperation rate on subject and match level.

Figure 3.4: Stability of Cooperation over Rounds



Notes: The lines show the average frequency of cooperation in the first 5 rounds of all and the last 3 supergames.

Figure 3.4 shows the average frequency of cooperation over the first five rounds. The cooperation rates start high (above 78 percent in all supergames, and above 82 percent in the last three supergames) and decline over rounds. In the last three supergames, the average cooperation rate has reduced by 19 percent in Cor and 9 percent in NoCor, which indicates more stability in NoCor. We regress cooperation on round index, treatment dummies and their interaction term to test whether the difference in stability is statistically significant. The result shows that the decline of cooperation is steeper in Cor than in NoCor ($p = 0.02$).

Result 2: Subjects are not more cooperative when correlated signals are introduced. We cannot reject the corresponding $H0$ and we conclude that correlated signals do not increase cooperation rate empirically. The decline of cooperation in the treatment with correlation is slightly steeper as compared to without signal correlation. Our strategy estimation shows that subjects from the Cor treatment do not fail to coordinate on the correlation structure of signals. The comparable frequency of cooperation between treatments is caused by the lower level of leniency in the chosen strategies.

3.5.3 Communication

Next, we take a closer look at the pre-play communication content and explore whether and in what way the average cooperation rate and strategy choices are influenced by the opportunity to communicate. Two research assistants classify subject-round communication observations into 39 sub-categories, with 5.67 classifications on average for each subject-round observation and 2313 classifications in total for the last three supergames. We merge the 39 sub-categories into four main categories: Coordination, Deliberation, Relationship and Trivia.[7] We determine Cohen's κ on the main categorical level.[8] The average κ is 0.35, indicating a fair level of agreement between the raters. Coordination is the effort to coordinate on future behavior. It includes explicit and implicit punishment threats when signals differ. According to Table 3.4, for both treatments, the category of Coordination has been covered in almost all pre-play chats

[7]Detailed description of the sub-categories and the mapping from sub-categories to main categories are in Table C.4.1 in Appendix C.4. We use the same classification scheme as Dvorak and Fehrler (2018) except two differences. First, to account for signal correlation, we add sub-categories regarding matching of signals. Second, because communication in our sessions has a pre-play structure, we eliminate their category "Information" and its sub-categories which regard the share of information in repeated communication.

[8]When the rating is random, agreement occurs with probability $p_i^2(Yes) + p_i^2(No)$, $i = 1, 2$, where $p_i(Yes)$ is the frequency of rater i classifying objects into any categories, and $p_i(No) = 1 - p_i(Yes)$.

in the last three supergames. Observations concerning the discussion of actions and strategies are classified into the Deliberation category. The category of Relationship includes all chats on promises, showing trust and distrust, pleas for trustworthy behavior. This category is the least frequent among the four. Another frequently observed category is Trivia. This category covers small talk, off-topic talks, and expressions of boredom.

Looking closely at sub-categorical levels (Table C.4.1 in Appendix C.4), from both treatments most subjects propose that they both play C. A small fraction of subjects in Cor talk about non-lenient GRIM punishment, while subjects in NoCor cannot agree on a plan for punishment. Interestingly, although subjects play CGRIM, topics including matching of the signals, implicit and explicit punishment when signals differ are rarely discussed. They appear below a frequency of 0.001.

Table 3.4: Frequency of Valid Codings per Individual-Round Observation

	Cor	NoCor
Coordination	0.99	0.99
Deliberation	0.37	0.26
Relationship	0.15	0.03
Trivia	0.89	0.96

Notes: Frequency of valid codings on main categorical level of subject-round observations. Data is from the last three supergames. A classification is considered to be valid if the classifications from the two raters agree. It is possible that an observation belongs to more than one category, resulting in both column sums to be larger than 1.

To answer the question how communication is related to cooperation, we conduct logistic regressions. The marginal effects are presented in Table 3.5, with the dependent variable being the first-round cooperation, and the explanatory variables being the dummies whether or not the pre-play communication falls into a main category. We control for the supergame indexes and socio-demographic and other subject-related characteristics. Data include all supergames, because the cooperation in round one is so high in the last three supergames that little variation exists. Standard errors are bootstrapped with 1000 repetitions and are two-way clustered on subject-match dimensions (Cameron et al., 2011). In the treatment Cor, Deliberation is positively correlated to the first-round cooperation. In treatment NoCor, a positive correlation is found for the category Trivia.

We further assign subjects to strategies on the basis of posterior probabilities, i.e. we map each individual to a strategy that has the maximum posterior probability based on SFEM. Among the 136 subjects, 49 use correlation-variational strategies (CGRIM or CTFT), 87 use non-correlation related strategies (ALLD, ALLC, TFT and WSLS). We look at the frequencies of main categorical communication regarding correlation-variational strategies users and users of other strategies. Figure 3.5 presents the result. Subjects who use correlation-related strategies have a similar communication pattern as the other subjects. There is only a small difference in Deliberation and Relationship ($p < 0.001$ for both categories). Subjects who are assigned as users of CGRIM or CTFT deliberate more, and are involved in talks to build up relationship more often.

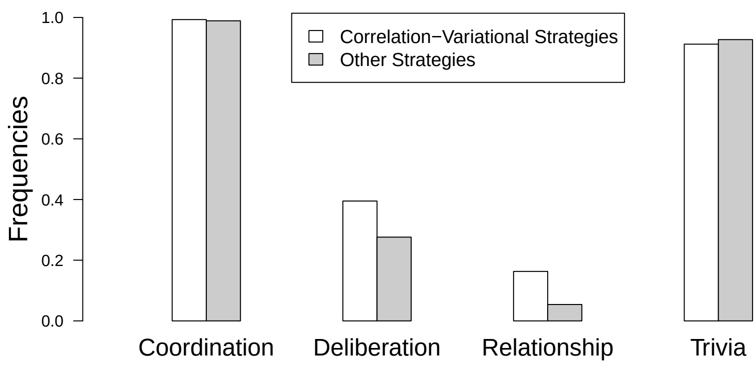
Table 3.5: Communication and First Round Cooperation

	Cor	NoCor
Coordination	0.96	1.38
Deliberation	0.65^{**}	0.09
Relationship	0.34	-0.21
Trivia	0.25	0.71^{*}
Supergame	0.30^{***}	0.11

Notes: The table presents the result of Logistic regression: the marginal effects of being in the main categories on first period cooperation. Data include all supergames. We control for socio-demographic and subject-related characteristics. Bootstrapped standard errors are calculated with two-way clustering, with cluster dimensions being subjects and match (1000 repetitions). Significance is based on the critical values of one-sided t statistics. $***$ ($**$,$*$) indicates significance on the 1 (5,10)% level.

Result 3: Almost all subjects talk about Coordination, attempting to coordinate their behavior. However only a small fraction of them in Cor discuss how to punish in case they deviate from the reward state. Trivia is another popular topic that almost everyone talks about. Although CGRIM attracts a substantial share of the likelihood in Cor, subjects do not threaten to punish neither implicitly nor explicitly when signals differ. In fact, they rarely mention to compare signals. The logistic regression of first round cooperation on communication categorical dummies shows that Deliberation and Trivia are positively related to the first-round cooperation of Cor and NoCor. The communication patterns are similar between correlation-variational strategy users and the other. There is only small difference in frequencies regarding

Figure 3.5: Frequency of Communication of Correlation-Variants and Other Strategies



Notes: Bars show the frequencies of communication of subjects who use correlation-variational strategies (CGRIM or CTFT) and other strategies. Individuals are assigned strategies with highest posterior probabilities. Data is from the last three supergames.

Deliberation and Relationship. Topics in these two categories are more frequently discussed by CGRIM and CTFT adopters.

3.6 Conclusion

Sustaining cooperation is difficult in the indefinitely repeated Prisoners' Dilemma when players cannot observe their opponents' actions perfectly. One way found in the literature to sustain cooperation is through correlated information (Awaya and Krishna, 2016). We investigate the effect of correlated information on cooperation in an indefinitely repeated Prisoners' Dilemma with public monitoring using laboratory experiments. The concept of correlated information is borrowed from Awaya and Krishna (2016).

In our study, information is correlated in a way such that when players choose the same actions, they receive identical information regarding the choice of the opponent. When their actions differ, information is independently determined by the opponent's action and an endogenously given noise level parameter. With correlation, a simple grim-trigger strategy exists that can be used efficiently support cooperation in equilibrium. This grim-trigger strategy cooperates as long as the two public signals exchanged in every round of the supergame match, and defects otherwise. Without correlation, no such strategy exists, and cooperation cannot be supported in equilibrium. Therefore, theory predicts more cooperation when signals are correlated. However, other studies have shown that people struggle to make correct inferences based on correlated information. Therefore, the hypothesis that correlation can foster cooperation by increasing the quality of imperfect monitoring is an empirical open question. We test the theoretical hypothesis in a lab experiment, which allows us to tightly control the correlation structure of public signals. The experimental data additionally allows us to answer the question if the correlation of signals is considered at the individual level. In addition, it is unclear whether subjects understand and use the correlated information in decision-making.

Our laboratory experiment consists of two between-subject treatments: Cor and NoCor. Signals are correlated in the former and are independently determined in the latter. To facilitate treatment comparison, we choose the stage game parameters such that cooperation is possible only in the Cor treatment. We find no treatment difference in average frequency of cooperation between treatments. Cooperation starts high in Cor, and has a steeper decline. An analysis of participants' strategies shows that a substantial fraction of participants of the Cor treatment take the correlation of the signals into account when making their choices. However, they do so by following a non-lenient punishment scheme for defection. As not all participants of the Cor treatment are cooperative, punishments happen frequently. This reveals two counteracting roles of correlation. On the one hand, correlation makes it possible to detect defection with certainty which makes efficient cooperation possible. On the other hand, the reaction to defection becomes non-lenient as there is no wiggle room surrounding bad signals.

Note that our implemented discount factor in the lab is very close to the threshold, which may explain why participants are extremely cooperative in the NoCor treatment. But it is still unclear why participants are surprisingly lenient and forgiving in the treatment without correlation. Dvorak and Fehrler (2018) also find considerable leniency under private monitoring with pre-play communication. This relates to the broader question why participants are so lenient towards bad signals under imperfect monitoring with pre-play. One potential explanation is that the pre-play communication allows them to connect with each other and social preferences may kick in.

Future research should investigate the effects of pre-play communication on cooperation and strategy choice. To gain insight into the limits of pre-play communication, it would be helpful to examine the play of the participants in the indefinitely repeated Prisoners' Dilemma with different PPE. Another potential agenda is to extend the design to imperfect private monitoring. The signals are private knowledge, and the participants report their received signals to their opponents. As subjects may not communicate truthfully, whether cooperation can still be sustained in such a scenario is an interesting question for future research.

Chapter 4

The Effect of Communication and Noise on Strategic Uncertainty in the Indefinitely Repeated Prisoners' Dilemma

Abstract

Sustaining cooperation is difficult in indefinitely repeated games when the partner's actions can only be observed with noise. Communication, especially before a repeated game, is an important determinant of cooperation. The literature proposes that communication sustains cooperation by reducing perceived strategic uncertainty. However, there is no direct empirical evidence on how subjects perceive strategic uncertainty when they can communicate with their partners. I study the effect of pre-play communication in an indefinitely repeated Prisoners' Dilemma using laboratory experiments. I vary the noise levels to control how accurately the subjects observe the action of their partners, and I elicit subjective beliefs about partners' actions to measure perceived strategic uncertainty. I find that pre-play communication helps sustain cooperation at all noise levels by reducing the perceived strategic uncertainty. Further elicitation tasks show that pre-play communication changes subjects' attitudes towards strategic uncertainty by making them more optimistic about receiving the higher expected outcome and more averse against uncertainty when the noise is low.

4.1 Introduction

In many repeated economic interactions, players cannot perfectly observe the past actions of other players. Instead, they observe noisy signals that represent these actions. The collusion model of oligopoly (Green and Porter, 1984), in which firms observe noisy market prices, is an example of imperfect public monitoring. Sustaining cooperation is difficult in indefinitely repeated games with imperfect monitoring. Pre-play communication, or communication before the game starts, has been shown to be an important determinant of cooperation in such scenarios (e.g., Dvorak and Fehrler 2018; Cooper and Kühn 2014; Embrey et al. 2013). Previous literature suggests that communication helps sustain cooperation by reducing strategic uncertainty (e.g., Kartal and Müller 2018; Dvorak and Fehrler 2018), meaning that players are more likely to believe that their partners will cooperate. However, there is no direct empirical evidence of whether communication changes the way subjects perceive uncertainty. It is also unclear whether the positive effect of communication on cooperation could carry over to scenarios in which cooperation is risky in the sense that the cost of cooperation exceeds its future gains. In this study, I present empirical evidence on how pre-play communication affects cooperation and perceived strategic uncertainty.

The aim of this study is to examine the effect of communication on cooperation in different scenarios, where a cooperative equilibrium may or may not be established. I answer two specific questions: 1) Does communication still sustain cooperation when cooperation is theoretically not possible? And does the effect of communication on cooperation differ in various scenarios, depending on whether a cooperative equilibrium can be established? 2) If communication affects cooperation, is it through reducing strategic uncertainty, or other channels?

To answer these research questions, I implement six treatments of indefinitely repeated Prisoners' Dilemma under public monitoring in the lab. With public monitoring, subjects do not receive feedback on their partner's real actions at the end of each round. Instead, they receive a noisy signal that incorrectly transmits the partner's action with a certain probability. Noisy signals are binary and can be either cooperative or defective. From the signals, subjects cannot infer what actions their partner has chosen. The round payment is determined by the player's own action and the signal of the other player's action. The treatments vary in two dimensions.

First, I vary the noise levels ϵ such that cooperative equilibria may only be established when the noise is not high. Higher noise makes it more difficult to receive the correct signal. Sustaining cooperation is more difficult when noise is high, because the probability of receiving a defective signal is high, even when the partner has played cooperatively. In this study, three noise levels are considered. The noise level is either low ($\epsilon = 0.05$), medium ($\epsilon = 0.20$), or high ($\epsilon = 0.35$). I choose these noise levels such that subgame-perfect equilibria (SPE) can be constructed with low and medium noise, but with a medium level of noise, cooperation is very risky for the subjects. Cooperative SPE does not exist with high noise.

The second dimension is communication. For each noise level, I implement one treatment with communication and one without. Communication takes place at the beginning of each interaction and is thus called pre-play communication. Communication is of the free form. Subjects write to their partners via a chat box for 120 seconds.

This 3×2 design allows me to examine whether communication sustains cooperation even when it is theoretically not feasible, and to what extent the effect of communication on cooperation differs across different noise treatments. Equilibrium concepts do not consider strategic uncertainty. Communication may affect cooperation by reducing strategic uncertainty. When communication is allowed, subjects are able to form prior knowledge regarding which actions or strategies their partners will play. This knowledge helps sustain cooperation by reducing the uncertainty associated with cooperation.

The definition of strategic uncertainty in repeated game literature is based on the risk dominance concept of Harsanyi and Selten (1988). In an indefinitely repeated Prisoners' Dilemma, a player may be uncertain about the choice that the other player will take. The equilibrium solution concepts overlook this uncertainty by assuming that subjects play grim-trigger. However, empirical studies, such as Dal Bó and Fréchette (2011); Blonski et al. (2011) show the importance of strategic uncertainty in decision-making. Subjects worry about the cost of being defected, and they cooperate more often when the cost is low. According to them, strategic uncertainty is defined as the amount of uncertainty associated with cooperation. In this paper, I elicit in each round the subjects' belief in the probability that their partner will cooperate in this round. If a subject believes that their partner will cooperate with a high probability, then according to the definition of strategic uncertainty, for this subject, there is low strategic uncertainty associated with cooperation. Pre-play communication may reduce the level of strategic uncertainty associated with cooperation, because subjects can coordinate and make promises before the game starts. Upon receiving a defective signal, it is more likely for them to attribute bad outcomes to noise, which may make them play strategies that are more lenient and forgiving instead of an inefficient grim-trigger.

Apart from affecting perceived strategic uncertainty towards cooperation, communication may also change subjects' attitudes towards strategic uncertainty itself. According to Baillon et al. (2018), a situation of uncertainty is an ambiguous situation in which subjects have knowledge only about the outcomes but not about the probabilities associated with these outcomes. Strategic uncertainty is a situation of ambiguity in which the uncertainty arises from strategic interactions, meaning that subjects are uncertain about what their partner is going to play. The indefinitely repeated Prisoners' Dilemma is a situation of strategic uncertainty, as the uncertainty arises from strategic interaction. In each round, each player faces ambiguity regarding the action their partner will take. This aversion towards strategic uncertainty is likely to make players hesitant to cooperate. I adopt the methodology of Bruttel et al. (2023) to measure subjects' attitudes towards strategic uncertainty. Specifically, I elicit the certainty equivalents for each of the two possible actions of the Prisoners' Dilemma in every round by asking subjects how much they would like to receive with certainty instead of participating in the game with this action. This method allows for breaking down uncertainty attitudes into two dimensions using a structural model. The first dimension measures an extra optimism or pessimism regarding the partner playing cooperation, which is modeled by an extra weight on the desirable expected outcome. The second dimension measures aversion against uncertainty, which is represented by adding a utility or dis-utility from the source of uncertainty to the model. In this paper, I investigate how noise and communication affect these two dimensions of uncertainty attitudes.

I find that pre-play communication increases cooperation rates across all noise levels, even in scenarios where cooperative equilibrium cannot be established. Communication makes strategies more lenient and reduces differences in the frequency of cooperation between treatments. Communication also reduces perceived strategic uncertainty towards cooperation, leading to higher and more accurate beliefs about cooperation. Additionally, communication affects attitudes towards strategic uncertainty in the low noise treatment, with a higher proportion of rounds in which subjects form optimistic beliefs about receiving the higher expected outcome and have higher aversion against uncertainty.

In the next section, I discuss the related literature. In section 3, I present the theoretical models. The experimental design is presented in section 4. In section 5, I discuss the research questions and main hypotheses. Section 6 presents the results.

4.2 Related Literature

This study contributes to the literature on imperfect public monitoring (Green and Porter, 1984) and related empirical studies (Aoyagi and Fréchette, 2009; Fudenberg et al., 2012; Embrey et al., 2017; Aoyagi et al., 2019). The formation of public monitoring in this study is the same as Aoyagi et al. (2019) but differs from Aoyagi and Fréchette (2009) and Fudenberg et al. (2012). The public signal in Aoyagi and Fréchette (2009) is continuous and depends on the sum of the actions of both players. In this study, there are two public signals that are binary, and each corresponds to the actions of both players. The round payoff in Fudenberg et al. (2012) is determined by both public signals, whereas in this study, the payoff is determined by its own action and the signal that represents the opponent's action.

Under imperfect monitoring, communication has been theoretically found to be an important factor that sustains cooperation (Compte, 1998; Obara, 2009; Awaya and Krishna, 2016). The empirical evidence from Dvorak and Fehrler (2018) confirms the importance of communication. They implement pre-play and repeated communication under imperfect public and private monitoring and find that both communication scenarios maintain cooperation. With pre-play communication, cooperation starts high but declines over rounds, whereas repeated communication stabilizes cooperation rates over rounds. They propose that communication affects cooperation by reducing strategic uncertainty; however, they do not directly measure the degree of strategic uncertainty perceived by subjects. Their main research focus is on the role of communication in different monitoring structures. The uncertainty in their treatments comes from differences in the monitoring structures, which are measured by the basin of attraction of defection (Dal Bó and Fréchette, 2011). While in this paper the uncertainty arises from different noise levels of the same monitoring structure.

Previous research has examined how noise levels affect cooperation and strategies under imperfect public monitoring. Empirical studies such as Fudenberg et al. (2012); Aoyagi et al. (2019) observe that under indefinitely repeated Prisoners' Dilemma, cooperation rates as well as the leniency of strategies decrease when noise increases. However, it is unclear why the strategies are more lenient with lower noise. It is also unclear whether these observations still occur in scenarios with pre-play communication.

One of the central questions here is with different noise levels, what beliefs subjects hold about which action their partner has chosen and will choose, or what strategies they believe their partner is playing. The belief that a player forms regarding the action of the other player is key to understanding equilibrium because players best respond to these beliefs. However, it is difficult to form accurate beliefs. Aoyagi et al. (2022) elicit beliefs in indefinitely repeated Prisoners' Dilemma with perfect monitoring without noise. They find that subjects who use different strategies form different beliefs that they tend to be overly optimistic that their partner is using similar strategies. It is even more challenging to form correct beliefs when the monitoring structure is imperfect because past actions are not observable, and strategies are more complex with multiplicity of equilibria. Similar to Aoyagi et al. (2022), this study elicits subjects' beliefs about their partner's action as a measure of the associated strategic uncertainty but with public monitoring. Unlike Aoyagi et al. (2022), this study does not estimate subjects' beliefs about supergame strategies because, with public monitoring, subjects do not receive feedback on their partner's actions, and the updating of belief is complicated. The elicited beliefs alone are sufficient to answer the research question.

This study is also related to the literature that addresses aversion to uncertainty (Camerer and Weber, 1992). Ambiguous situations with unknown probabilities are preferred less frequently than those with known probabilities. Strategic uncertainty is a type of ambiguous situation in which uncertainty arises from interactions between players. Various experiments have been conducted to identify aversion to strategic uncertainty (e.g., see Heinemann et al. 2009; Kelsey and le Roux 2014; Greiner 2023; Bruttel et al. 2023). This study uses the same method as Bruttel et al. (2023) to measure the subjects' general attitudes toward situations of strategic uncertainty and to identify their optimism or pessimism regarding certain outcomes. While past works have focused on one-shot games without communication, this study applies their methodology to repeated interactions with communication.

4.3 Theoretical Models

This section first introduces the repeated Prisoners' Dilemma with public monitoring and the perfect-public equilibrium solution concept. It then introduces a theoretical framework for identifying certainty attitudes.

4.3.1 Repeated Prisoners' Dilemma of Public Monitoring

Two players infinitely repeat a 2×2 stage game. In each round, both players simultaneously choose an action from the action set. The action set for player i is $A_i = \{C, D\}$, which means that player i can either cooperate or defect. An action is transmitted into a signal with a transmission error ϵ. The signal set of player i's action is given by $\Omega_i = \{c, d\}$, which means that the signal can either be cooperative or defective. I denote the realized signal that player i receives regarding player j's action as ω_j. For each action profile $a = (a_i, a_j)$, a conditional probability distribution $p(\omega|a)$ is assigned to the signals. Specifically, $p(\omega_i = c|a_j = C) = p(\omega_i = d|a_j = D) = 1 - \epsilon$ and $p(\omega_i = c|a_j = D) = p(\omega_i = d|a_j = C) = \epsilon$. At the end of each round, the signal profile $\omega = (\omega_i, \omega_j)$ is observed by both players.

The payoff for player i is determined by the player's action a_i and the signal ω_j regarding player j's action. Denote the stage game payoff of player i as $g_i(a_i, \omega_j)$. The expected payoff from the stage game is $u_i(a) = \sum_{\omega_j \in \Omega_j} g_i(a_i, \omega_j) p(\omega_j|a)$.

The expected stage game payoff (u_i, u_j) can be normalized and represented as the normal form as follows:

	C	D
C	$1, 1$	$-l, 1+g$
D	$1+g, -l$	$0, 0$

For this game to be Prisoners' Dilemma type of game, the sucker payoff $-l$ should be smaller than the punishment payoff 0, and the temptation payoff $1 + g$ should be larger than the temptation payoff and larger than the reward payoff 1: $l > 0$, $g > 0$ and $g < 1 + l$.

Let $\Omega^t = \{\Omega_i, \Omega_j\}^t$ be a set of public histories up to round t. A public strategy for player i is a mapping $\sigma_i : \bigcup_{t \geq 0} \Omega^t \rightarrow \Delta A_i$. A strategy profile $\sigma = (\sigma_i, \sigma_j)$ is a perfect-public equilibrium (PPE) if σ_i is a public strategy, and for any history up to round t, σ is a Nash equilibrium in all rounds following t. PPE is a subgame perfect Nash equilibrium (SPE) that depends only on public signals.

Let δ be the discount factor for both players. Cooperation can be sustained by non-lenient and non-forgiving strategies, such as the grim-trigger strategy. Under public monitoring, the SPE depends on public signals. The grim-trigger starts with cooperation and punishment is triggered upon receiving a defective signal. Punishment is non-forgiving. Once players enter the punishment state, they will never revert to cooperation. This grim-trigger strategy under public monitoring is PPE if the long-run benefit from cooperation $\frac{1}{1-\delta(1-\epsilon)^2}$ is at least as large as the gain from defecting $\frac{1+g}{1-\delta\epsilon(1-\epsilon)}$. The common discount factor δ should thus be sufficiently large

$$\delta \geq \delta^{PPE} = \frac{g}{(1+g)(1-\epsilon)^2 + \epsilon^2 - \epsilon} \tag{4.1}$$

4.3.2 Strategic Uncertainty

PPE is not sufficient to understand subjects' behavior in the lab because it overlooks strategic uncertainty by assuming that players are sure that the other will follow the grim-trigger strategy (Dal Bó and Fréchette, 2018). This is problematic, as it assumes that strategic uncertainty and aversion to strategic uncertainty do not impact behavior. Uncertainty about what strategy the partner is playing causes the player to worry about getting the sucker payoff $-l$. $-l$ affects cooperation but the PPE condition does not depend on $-l$.

To address this problem, Dal Bó and Fréchette (2011) propose a condition based on the risk dominance concept of Harsanyi and Selten (1988). Suppose that a player plays either the cooperative grim-trigger (GRIM) or the always-defect (ALLD). They define a threshold basin of attraction of defection (BAD). This threshold is the probability that the partner plays GRIM, such that the player is indifferent between GRIM and ALLD. Denote π as the probability that the partner plays GRIM. The expected payoff from playing GRIM is $\frac{\pi}{1-\delta} - (1-\pi)l$ and the expected payoff from playing ALLD is $\pi(1+g)$. This player is indifferent between GRIM and ALLD if $\pi = \frac{l}{l-g+\frac{\delta}{1-\delta}}$.

This threshold can be extended to the public monitoring structure. For GRIM under public monitoring, punishment is triggered if at least one of the public signals is defective. The expected payoff from playing GRIM is $\pi\frac{1}{1-\delta(1-\epsilon)^2} - (1-\pi)\frac{l}{1-\delta\epsilon(1-\epsilon)}$. The expected payoff from playing ALLD is $\pi\frac{1+g}{1-\delta\epsilon(1-\epsilon)}$. The BAD threshold that makes a player indifferent between GRIM and ALLD is thus:

$$\pi_{BAD} = \frac{l}{l - g + \frac{\delta((1-\epsilon)^2-\epsilon(1-\epsilon))}{1-\delta(1-\epsilon)^2}} \tag{4.2}$$

The size of BAD is a measure of how robust cooperation is under strategic uncertainty. Under perfect monitoring, cooperation rates increase as the size of BAD decreases (Dal Bó and Fréchette, 2018). If the size of BAD is high, players need to believe that their partner will play cooperative GRIM with high probability to play cooperatively.

4.3.3 Identification of Attitudes Towards Strategic Uncertainty

This section follows from the theoretical model in Bruttel et al. (2023) which elicits aversion to strategic uncertainty. The stage game of Prisoners' Dilemma can be interpreted as having two lotteries. Lottery C has two outcomes, $g(C,c)$ and $g(C,d)$. Lottery D has the outcomes $g(D,c)$ and $g(D,d)$. The probabilities associated with these outcomes are unknown because of strategic uncertainty. The value of a lottery for player i is:

$$W_i(a_i, \omega_j, \pi_i) = E[u_i(g(a_i, \omega_j))|\pi_i] + \Delta_i[g(a_i, \omega_j)|\pi_i] \tag{4.3}$$

Where $u(.)$ is the utility function, $g(a_i, \omega_j)$ are possible outcomes, and π_i is player i's subjective probability distribution over the outcomes. Δ_i is the player's attitude towards strategic uncertainty. If the player is a rational expected utility maximizer, $\Delta_i[g(a_i, \omega_j)|\pi_i] = 0$; otherwise it is non-zero. The player's uncertainty attitudes has two dimensions: optimism and pessimism towards the desirable expected outcome, and preference for or aversion against the source of uncertainty. Δ_i can thus modeled using two parameters, α_i and λ_i. Each parameter represents one dimension of the attitudes. Equation 4.3 can be rewritten using α_i and λ_i:

$$\begin{aligned} W_i(a_i, \omega_j, \pi_i) = (\pi_i + \alpha_i)u_i(&\sum_{\omega_j \in \Omega_j} g_i(a_i, \omega_j)p(\omega_j|C)) \\ &+ (1 - \pi_i - \alpha_i)u_i(\sum_{\omega_j \in \Omega_j} g_i(a_i, \omega_j)p(\omega_j|D)) - \lambda_i \end{aligned} \tag{4.4}$$

Player i believes that the other player will cooperate with probability π_i. The probability that the other player will defect is thus $1 - \pi_i$. The expected stage game payoff conditioning on subjective belief is $\pi_i u_i(\sum_{\omega_j \in \Omega_j} g_i(a_i, \omega_j)p(\omega_j|C)) + (1 - \pi_i)u_i(\sum_{\omega_j \in \Omega_j} g_i(a_i, \omega_j)p(\omega_j|D))$. Parameter α_i captures how optimistic the player is towards partner playing cooperation. It is a subjective weight that the player puts on the expected outcome that describes how optimistic the player is towards receiving the expected payoff $\sum_{\omega_j \in \Omega_j} g_i(a_i, \omega_j)p(\omega_j|C))$ rather than

$\sum_{\omega_j \in \Omega_j} g_i(a_i, \omega_j)p(\omega_j|D)$. If $\alpha_i > 0$, the player is optimistic. If $\alpha_i < 0$, the player is pessimistic. If the player is neither optimistic nor pessimistic, but has a pure Bayesian view, $\alpha_i = 0$. Notice that α us not bounded between $[-1, 1]$. The parameter λ_i is the player's preference for or aversion against strategic uncertainty. It is modeled as an utility or dis-utility from uncertainty. If $\lambda_i > 0$, the player is averse to strategic uncertainty because λ_i reduces the value of the lottery. If $\lambda_i < 0$, the player prefers strategic uncertainty. $\lambda_i = 0$ means the player is neutral towards strategic uncertainty.

In a 2×2 game, parameters α_i and λ_i can be solved mathematically using a system of equations. In the Prisoners' Dilemma, players evaluate both lottery C and lottery D. The values of these two lotteries are:

$$\begin{aligned} W_i(C, \omega_j, \pi_i) = (\pi_i + \alpha_i)u_i(\sum_{\omega_j \in \Omega_j} g_i(C, \omega_j)p(\omega_j|C)) \\ + (1 - \pi_i - \alpha_i)u_i(\sum_{\omega_j \in \Omega_j} g_i(C, \omega_j)p(\omega_j|D)) - \lambda_i \\ = Eu_i(g_i(C, \omega_j)|\pi_i, p(\omega_j|a_j)) \\ + \alpha_i(u_i(\sum_{\omega_j \in \Omega_j} g_i(C, \omega_j)p(\omega_j|C)) - u_i(\sum_{\omega_j \in \Omega_j} g_i(C, \omega_j)p(\omega_j|D))) - \lambda_i \end{aligned} \tag{4.5}$$

$$\begin{aligned} W_i(D, \omega_j, \pi_i) = (\pi_i + \alpha_i)u_i(\sum_{\omega_j \in \Omega_j} g_i(D, \omega_j)p(\omega_j|C)) \\ + (1 - \pi_i - \alpha_i)u_i(\sum_{\omega_j \in \Omega_j} g_i(D, \omega_j)p(\omega_j|D)) - \lambda_i \\ = Eu_i(g_i(D, \omega_j)|\pi_i, p(\omega_j|a_j)) \\ + \alpha_i(u_i(\sum_{\omega_j \in \Omega_j} g_i(D, \omega_j)p(\omega_j|C)) - u_i(\sum_{\omega_j \in \Omega_j} g_i(D, \omega_j)p(\omega_j|D))) - \lambda_i \end{aligned} \tag{4.6}$$

Assume that the players have a CRRA utility function with r_i as the relative risk aversion parameter. Players' willingness to accept (WTA) the lottery is a direct measure of their evaluation of the lottery. The WTA can be elicited via the Becker-DeGroot-Marschak (BDM) mechanism (Becker et al., 1964). Equation 4.5 and 4.6 can be written as

$$\lambda_i = \frac{(WTA_i(C,\omega_j,\pi_i))^{1-r_i} - (Eu_i(g_i(C,\omega_j)|\pi_i, p(\omega_j|a_j)))^{1-r_i}}{1-r_i} + \frac{\alpha_i((\sum_{\omega_j\in\Omega_j} g_i(C,\omega_j)p(\omega_j|C))^{1-r_i} - (\sum_{\omega_j\in\Omega_j} g_i(C,\omega_j)p(\omega_j|D))^{1-r_i})}{1-r_i} \tag{4.7}$$

$$\lambda_i = \frac{(WTA_i(D,\omega_j,\pi_i))^{1-r_i} - (Eu_i(g_i(D,\omega_j)|\pi_i, p(\omega_j|a_j)))^{1-r_i}}{1-r_i} + \frac{\alpha_i((\sum_{\omega_j\in\Omega_j} g_i(D,\omega_j)p(\omega_j|C))^{1-r_i} - (\sum_{\omega_j\in\Omega_j} g_i(D,\omega_j)p(\omega_j|D))^{1-r_i})}{1-r_i} \tag{4.8}$$

Solving equations 4.7 and 4.8 yields:

$$\alpha_i = \frac{A}{B} \tag{4.9}$$

where

$$A = (WTA_i(D,\omega_j,\pi_i))^{1-r_i} - (WTA_i(C,\omega_j,\pi_i))^{1-r_i} + Eu_i(g_i(C,\omega_j)|\pi_i, p(\omega_j|a_j)))^{1-r_i} - Eu_i(g_i(D,\omega_j)|\pi_i, p(\omega_j|a_j)))^{1-r_i} \tag{4.10}$$

$$B = (\sum_{\omega_j\in\Omega_j} g_i(C,\omega_j)p(\omega_j|C))^{1-r_i} - (\sum_{\omega_j\in\Omega_j} g_i(C,\omega_j)p(\omega_j|D))^{1-r_i} - (\sum_{\omega_j\in\Omega_j} g_i(D,\omega_j)p(\omega_j|C))^{1-r_i} + (\sum_{\omega_j\in\Omega_j} g_i(D,\omega_j)p(\omega_j|D))^{1-r_i} \tag{4.11}$$

λ_i can be solved by plugging α_i into equation 4.7.

4.4 Experimental Design

I implement an indefinitely repeated Prisoner's Dilemma under public monitoring. The treatments vary in the level of noise and whether pre-play communication is allowed. All the treatments are implemented under imperfect public monitoring. Two players choose actions $a_i \in \{C, D\}$ simultaneously in each round. At the end of each round, subjects do not observe real actions but receive noisy signals representing the real actions. Specifically, they are informed of $\{a_i, \omega_i, \omega_j\}$, where ω_i and ω_j denote the signals representing one's own action and the partner's action, respectively.

There are 8 subjects in each matching group and 6 matching groups in each treatment.[1] An infinitely repeated Prisoners' Dilemma is implemented in the lab as an indefinitely repeated interaction with continuation probability δ after every one. In this study, $\delta = 0.8$ for all treatments. A supergame is an indefinitely repeated interaction. In each supergame, subjects are matched to a new partner within their matching group so that they do not encounter the same partner twice. The subjects play six supergames in total.

The lengths of the supergames are predetermined. Specifically, I generate two sequences of random numbers and use them to determine the length L_i of supergame i[2]. Each sequence is used repeatedly for three matching groups for each treatment. To increase the number of observations, an adapted version of the block-random-termination method (Fréchette and Yuksel, 2017) is used. Subjects play a block of five rounds and are informed at the end of the fifth round whether the supergame ends. If the supergame ends, they are told in which round it ends. Otherwise, the supergame continues and the subjects are informed about whether the supergame ends after every round.

4.4.1 Treatments

There are 6 between-subject treatments that vary in two dimensions.

[1]This sample size gives me enough power (88.30%) to detect a effect size of 10% with one-sided t-test.

[2]I generate two sequences of uniformly distributed random variables between 0 and 1. The random seeds of the two sequences are 10 and 11. Seeds 1 to 9 have been used in my other research projects, and thus, I start from 10. The two sequences are denoted by $\{r_n\}_i = \{r_1, r_2, ..., r_x\}_i$, where $i = 10, 11$ indicates the seed underlying the sequence, and $n \in \mathbb{N}$. The first supergame has x_1 rounds, if $r_{x_1} \leq 0.2$ and for all $n < x_1$, $r_n > 0.2$. The second supergame has $x_2 - x_1$ rounds if $r_{x_2} \leq 0.2$, for all $x_1 < n < x_2$, $r_n > 0.2$, and so forth. The resulting (lengths of) sequences are SQ1 $(5, 6, 3, 10, 5, 6)$ and SQ2 $(1, 4, 7, 6, 10, 1)$.

TN05: The noise is low ($\epsilon = 0.05$). The subjects do not engage in pre-play communication.

TP05: The noise is low ($\epsilon = 0.05$). Subjects communicate with their partner via a chat box at the beginning of each supergame.

TN20: The noise is intermediate ($\epsilon = 0.20$). The subjects do not engage in pre-play communication.

TP20: The noise is intermediate ($\epsilon = 0.20$). Subjects communicate with their partner via a chat box at the beginning of each supergame.

TN35: The noise is high ($\epsilon = 0.35$). The subjects do not engage in pre-play communication.

TP35: The noise is high ($\epsilon = 0.35$). Subjects communicate with their partner via a chat box at the beginning of each supergame.

4.4.2 Experimental Procedure

All the treatments consist of three parts. In the first three supergames, subjects play the indefinitely repeated Prisoners' Dilemma under public monitoring. In the last three supergames, apart from the indefinitely repeated Prisoners' Dilemma, subjects answer three additional questions in each round that elicit their WTAs of both lotteries and their belief about the probability that the partner will cooperate in this round. At the end of the experiment, subjects state the WTAs for 22 lotteries, which are used to estimate their risk preference parameter r_i. The detailed experimental procedure is as follows:

Instructions. The subjects read the written instructions and answer the control questions. Examples of instructions and control questions can be found in the Appendix.

Communication. In treatments with communication, subjects begin each supergame with a pre-play communication stage of 2 minutes. They are allowed to freely chat with their partner via a chat box. In treatments without communication, subjects start directly with Prisoners' Dilemma.

Stage game. In each round, the subjects choose between two actions $a_i \in \{C, D\}$ simultaneously. At the end of each round, they receive public signals regarding their chosen actions.

Belief and certainty equivalents elicitation. In the last three supergames, apart from choosing actions in each round, the subjects are asked to report their belief about their partner's action in this round. Furthermore, they answer two questions which elicit their certainty equivalents of the stage game.

Risk preference task. When all supergames are finished, subjects face lotteries of two outcomes under 11 probabilities ranging from 0 to 100. I elicit their certainty equivalents for each of these lotteries.

Questionnaire. At the end of the experiment, subjects answer a questionnaire that has two parts. The first part of the questionnaire is a a socio-demographic survey. Questions in the second part ask subjects to look back at their decision-making and answer questions such as how close they feel to their partners overall, what motivates their decision-making and so on. A complete questionnaire of the second part can be found in the Appendix D.7.

4.4.3 Parameterization of the Stage Game

The stage game payoff matrix is presented in Table 4.1. The payoffs are in experimental currency units. Figure 1 shows the expected stage game payoffs for all three noise levels: $\epsilon = 0.05$, $\epsilon = 0.20$, $\epsilon = 0.35$. The normalized expected stage game parameters are $g \approx 0.35$ and $l \approx 0.46$ for $\epsilon = 0.05$, $g \approx 0.70$ and $l \approx 0.83$ for $\epsilon = 0.20$, and $g \approx 6.09$ and $l \approx 6.64$ for $\epsilon = 0.35$. The stage game parameters ensure that cooperative equilibria can only be constructed when $\epsilon = 0.05$ and $\epsilon = 0.20$ and that cooperation is risky with $\epsilon = 0.20$ and $\epsilon = 0.35$ in terms of BAD. Table 4.2 summarizes the PPE and BAD thresholds for the three noise levels. The aim of this design is to study whether subjects still cooperate when it is theoretically not a good option. TP35 has very high noise, which makes cooperation only slightly more profitable than defection. Whether pre-play communication still supports cooperation in such an extreme case is an interesting question.

Table 4.1: Stage game Payoff

	c	d
C	32	4
D	38	12

Table 4.2: Summary of PPE and BAD

	$\epsilon = 0.05$	$\epsilon = 0.20$	$\epsilon = 0.35$
SPE	0.30	0.75	-
BAD	0.18	0.97	-

Note: The perfect-public equilibrium and the basin of attraction of defection threshold do not exist at $\epsilon = 0.35$.

Figure 4.1: Expected Stage game Payoffs

	C	D
C	$30.6, 30.6$	$5.4, 36.7$
D	$36.7, 5.4$	$13.3, 13.3$

	C	D
C	$26.4, 26.4$	$9.6, 32.8$
D	$32.8, 9.6$	$17.2, 17.2$

	C	D
C	$22.2, 22.2$	$13.8, 28.9$
D	$28.9, 13.8$	$21.1, 21.1$

Notes: From top to bottom, expected payoffs are calculated with noise level 0.05, 0.20 and 0.35 respectively.

4.4.4 Elicitation Tasks

To estimate subjects' attitudes towards strategic uncertainty, I elicit their subjective beliefs about their partner's action and their certainty equivalents in the stage game. The elicitation method follows that of Bruttel et al. (2023). The main concern when adding an elicitation stage between rounds is that the tasks may alter the subjects' behavior. To address this issue, elicitation starts only from the fourth supergame.[3]

[3]This separation is to address the concern that asking for subjects' belief and WTAs may alter their mindset and their strategies. In the Appendix D.1, I compare the average frequency of cooperation in early and late supergames as a robustness check. The cooperation rates differ between early and late supergames in TP05, and are similar in the other treatments.

Belief elicitation. In every round, after subjects have decided on their action, they are asked to report the probability that they believe their partner cooperates in this round. The specific question is as follows:

> *Question 1:* How likely do you think that the other player will choose option C in this round?

Subjects enter an integer between 0 and 100 as their probability estimation. This task is incentivized through binarized scoring rule. Following Danz et al. (2022), I present the quantitative details on incentives only on demand to avoid false reports and pull-to-center.[4]

Certainty equivalents for stage game lotteries. I elicit two certainty equivalents in each round. Subjects are asked to state how much they would like to receive in order to opt out from the stage game. The specific questions are:

> *Question 2:* Suppose the computer decides for you in this round and chooses option C. How many points would you like to have at least in order not to proceed with option C?

> *Question 3:* Suppose the computer decides for you in this round and chooses option D. How many points would you like to have at least in order not to proceed with option D?

Subjects state a value from the lowest possible payoff to the highest possible payoff in the stage game payoff matrix. According to Table 4.1, their report should be equal to or above 4 or 12, and should not exceed 32 or 38 for lottery C and lottery D respectively.[5] The computer randomly chooses one out of the last three supergames and then one round from that supergame to pay subjects based on their certainty equivalents of either lottery C or lottery D. The chances of both lotteries being chosen are equal. Subjects do not know which supergame is selected until the end of the experiment.

The certainty equivalent based payment is implemented using the BDM mechanism. The computer draws a random number between the minimum and the maximum stage game payoffs. If the stated value is less than or equal to the random number, the subject is paid with this random number. If the stated value is larger than the random number, the subject's payment

[4] Only one subject asked for the quantitative details of the elicitation method at the end of one session.

[5] From behavioral perspective, subjects' WTAs may go beyond or below the payoff limit if they derive utility from interacting with another player or the other way round. This study does not take such extreme attitudes into account.

is determined by the action the computer chooses for them and the signal representing their partner's action.

At the end of each round, subjects receive feedback on their own action, the public signals representing their own and their partner's action. In addition, they receive feedback on the random numbers for lotteries C and D, and the number of points they would earn if lottery C or lottery D of this round were chosen.

Risk attitude elicitation. I elicit subjects' risk preferences by asking them to state their WTAs for a list of lotteries. There are 22 lotteries for the two sets of outcomes. The values of the outcomes are the same as the stage game payoff. Subjects receive either 4 or 32 or either 12 or 38. The probability of realizing the higher outcome ranges from 0 to 100 in steps of 10.

Subjects specify their WTA for each lottery by answering the following question:

> *Question 4:* If you were allowed to choose between getting a certain number of points safely or playing the lottery, what is the minimum number of points you would like to opt out of the lottery?

The computer randomly chooses one lottery and then a random number is drawn for that lottery. If the random number is greater than or equal to the subject's stated value, the subject is paid with the random number. Otherwise, they are paid according to the lottery.

4.4.5 Payment

The payment is determined in three parts. In supergames one to three, the stage game payoffs of all payoff-relevant rounds count towards the final payment. To avoid hedging, one of the last three supergames is paid with the stage game payoff, one is paid with the belief elicitation task, and the remaining one is paid with the certainty equivalence elicitation task. The supergames are chosen randomly by computer. Subjects do not know which supergame is paid with which task. If the supergame is paid with the stage game payoffs, the points from all payoff-relevant rounds are counted. If the supergame is paid with the belief elicitation task, the computer randomly selects one round. The payment is determined using the binarized scoring rule. Two random numbers between 0 and 100 are drawn. If the stated belief estimation is larger than any of the two random numbers, the subject receives 50 points. If the supergame is paid with the certainty equivalence task of the lotteries, then the computer randomly selects one round. One of the two lotteries is paid. The third part of the payment is the risk attitude elicitation task. One of the lotteries is paid.

4.4.6 Sessions

All treatments are programmed in zTree (Fischbacher, 2007). The sessions are conducted in 2023. Two sessions of four matching groups are conducted in the WISO lab of Hamburg University, and the remaining sessions are conducted in Brelab of Bremen University. Table 4.3 summarizes these sessions.[6] Subjects are students from both universities, and they are recruited via hroot in Hamburg (Bock et al., 2014), and via ORSEE in Bremen (Greiner, 2004). Each session takes less than two hours with an average payment of around 25 euros.

Table 4.3: Summary of the experimental sessions

	TP05	TP20	TP35	TN05	TN20	TN35
sessions	4	4	3	3	2	1
matching groups	4	4	3	3	3	2
subject numbers	32	32	24	24	24	16
place	Bremen	Bremen	Bremen	Bremen	Bremen, Hamburg	Hamburg

4.5 Research Questions and Hypotheses

There are two main research questions. As a first research question, I investigate how noise levels interact with communication and what effects they have on cooperation rates. Second, I answer whether communication affects cooperation by reducing strategic uncertainty. The specific research questions are as follows.

Question 1. At a given noise level, how does pre-play communication affect cooperation?

H1.1: Subjects in the pre-play communication treatments are more cooperative than those in the no-communication treatments.

[6] At the time this paper is written, only half of the data has been collected. Data for all pre-play communication treatments and TN05 is collected in Bremen. For TN20, I conduct one session with one matching group in Bremen, and one session with two matching groups in Hamburg. For TN35, I conduct one session with two matching groups in Hamburg.

Dvorak and Fehrler (2018) has found that at a given noise level, pre-play communication increases cooperation rates. Their noise level $\epsilon = 0.1$ yields a PPE threshold of 0.65, and the size of BAD is 0.75. These values are similar to those in the treatment with $\epsilon = 0.20$ in this study. It is unclear whether the results will continue to carry over to high-noise treatment.

Question 2. How does noise affect cooperation rates in a given communication scenario?

H2.1: Subjects are more cooperative in treatment TN05 than in TN20, and more cooperative in treatment TN20 than in TN35. They are more cooperative in treatment TP05 than in TP20 and more cooperative in TP20 than TP35.

Fudenberg et al. (2012); Aoyagi et al. (2019) look at cooperation in different noise treatments without communication and find that without communication, cooperation rates decrease as noise increases. There is no theoretical and empirical evidence so far on the cooperation behavior at different noise levels when communication is possible. According to the PPE thresholds in Table 4.2, it is likely that with pre-play communication, the average cooperation rate is higher in the lower noise treatment than in the higher noise treatment. Whether the effect of communication decays more quickly in the high-noise treatment than in the lower-noise treatment is unclear, and this paper will try to answer this question. This paper will also examine the difference in cooperation rates among the pre-play communication treatments and compare it to the difference in cooperation rates among the no-communication treatments. I expect that communication reduces the treatment difference in cooperation by improving the cooperation rates in TP35, so that the difference of the former is smaller than the latter.

Question 3. Is there a difference in perceived strategic uncertainty associated with cooperation?

H3.1: At a given noise level, the subjective belief about the probability that the other player will cooperate is higher in treatments with pre-play communication than in those without.

H3.2: Under the same communication scenario, the subjective belief about the probability that the other player will cooperate is higher with $\epsilon = 0.05$ than with $\epsilon = 0.20$ and higher with $\epsilon = 0.20$ than with $\epsilon = 0.35$.

Subjects are more uncertain about their partner's action when the noise is high. Communication reduces strategic uncertainty regarding cooperation because subjects can coordinate strategies and make promises. I will also look at how beliefs change over the round depending on the received signals, which gives me information about how subjects interpret the signal.

The following questions are exploratory.

Question 4. Is there a difference in attitudes towards strategic uncertainty between treatments with and without pre-play communication? Is there difference in attitudes towards strategic uncertainty between treatments in the same communication scenario but with different noise levels?

Similar to Bruttel et al. (2023), the estimation of aversion towards strategic uncertainty follows a two-step procedure. First, I estimate subjects' relative risk aversion parameter r using the method of Hey et al. (2009). Second, I calculate α and λ of each subject in every round. I examine how their optimism and ambiguity aversion attitudes change over the rounds. I expect that subjects in the pre-play communication treatments are more optimistic about their beliefs and they remain optimistic after receiving a bad signal.

Question 5. What strategies do subjects play?

To answer this question, I estimate the strategies that subjects play based on the strategy frequency estimation method of Dal Bó and Fréchette (2011). I use the R package stratEst (Dvorak, 2018). This method obtains the maximum-likelihood estimates of the shares of a candidate strategy set. This candidate set has 20 strategies taken from Fudenberg et al. (2012). Additional I include three behavior strategies that are motivated by Backhaus and Breitmoser (2018) for robustness check.

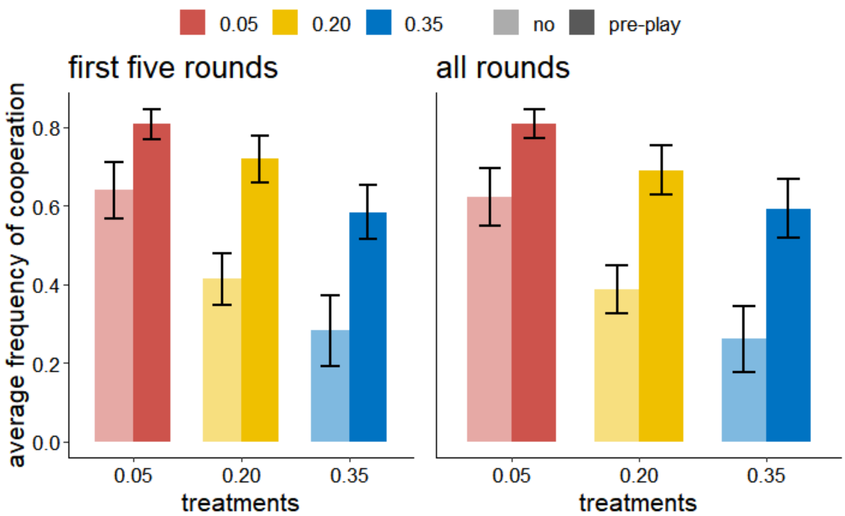
4.6 Results

To account for learning, supergame one is excluded from the analysis. As the belief elicitation task and certainty equivalence task are introduced in supergame four, supergame four is also excluded. The results focus mainly on the first five rounds of each supergame, so that the transition from block termination to random termination after the fifth round does not confound the results.

4.6.1 Cooperation Rates

Figure 4.2 shows the average cooperation rates for all the treatments. Pre-play communication increases the cooperation rates for treatments at all noise levels ($p_{0.05} = .015$, $p_{0.20} < .001$, $p_{0.35} = .006$). The effect is 16% in the low noise $\epsilon = 0.05$ treatments, and 30% with noise $\epsilon = 0.20$ and $\epsilon = 0.35$. Without pre-play communication, the average cooperation rates between TN20 and TN35 are both low and do not differ significantly ($p_{TN20vs.TN35} = .122$). Subjects cooperate more often in TN05 than in TN20 ($p_{TN05vs.TN20} = .010$), and more often in TN05 than in TN35 ($p_{TN05vs.TN35} = .002$). With pre-play communication, subjects in TP20 are as cooperative as subjects in TP35 ($p_{TP20vs.TP35} = .069$), because communication improves the mean frequency by the same percentage for both treatments. The difference between TP05 and TP20 goes away ($p_{TP05vs.TP20} = .099$), and the cooperation rates are higher in TP05 than in TP35 ($p_{TP05vs.TP35} = .002$). Difference-in-differences tests show that the effect of pre-play communication is similar at all noise levels ($p_{0.05vs.0.20} = .189$, $p_{0.20vs.0.35} = .489$, $p_{0.05vs.0.35} = .233$).

Figure 4.2: Average frequency of cooperation

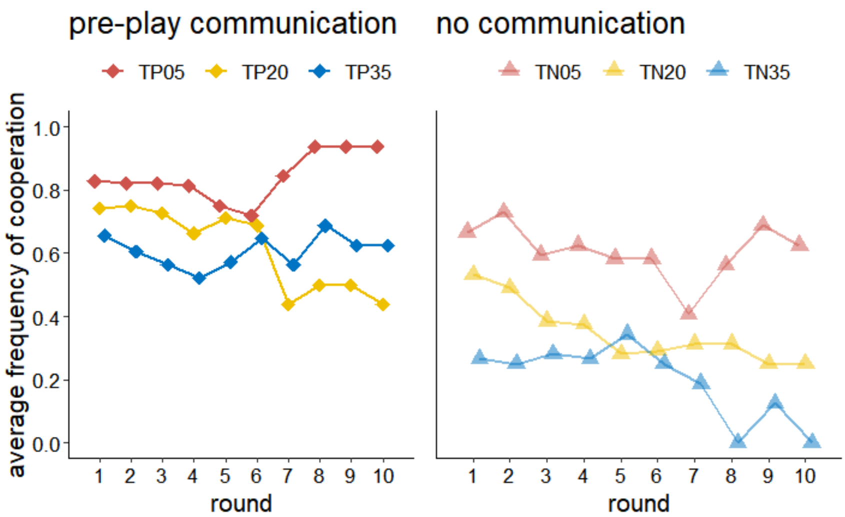


Note: This figure shows the average rate of cooperation of all the treatments. Data of all the rounds are included in the right panel. The left panel uses data of the first five rounds. The error bars are the standard error which are two-way clustered on subject level and match level.

Although cooperation is risky when the noise is 0.20 or 0.35, the average cooperation rates are high when subjects engage in pre-play communication. Pre-play communication sustains cooperation even when cooperation is risky and when a cooperative equilibrium cannot be constructed.

Similar results are obtained in the first round. The average frequency of cooperation in round one is shown in Figure 4.3. Pre-play communication increases the cooperation rates in the initial round at all noise levels ($p_{0.05} = .027$, $p_{0.20} = .023$, $p_{0.35} = .002$). With pre-play communication, subjects are more cooperative in TP05 than in TP35 ($p_{TP05vs.TP35} = .016$), but cooperation rates are similar between TP05 and TP20 ($p_{TP05vs.TP20} = .126$), and between TP20 and TP35 ($p_{TP20vs.TP35} = .189$). Without pre-play communocation, the differences in cooperation rates between treatments are larger. The average cooperation rates are higher at lower noise levels ($p_{TN05vs.TN20} = .118$, $p_{TN20vs.TN35} = .029$, $p_{TN05vs.TN35} = .002$). Pre-play communication reduces the differences in first-round cooperation. Difference in differences tests show that the effect is similar in the initial round at all noise levels ($p_{0.05vs.0.20} = .444$, $p_{0.20vs.0.35} = .162$, $p_{0.05vs.0.35} = .128$).

Figure 4.3: Average frequency of cooperation over rounds



Note: This figure shows the changes in average cooperation rates over the round. The left panel shows the results of the treatments with pre-play communication. The right panel shows the results for treatments without communication.

Figure 4.3 shows the changes in cooperation over the rounds. The average cooperation rates decline by approximately 5% in TP05 and TP35 in the first five rounds ($p_{TP05} = .025$, $p_{TP20} = .066$, $p_{TP35} = .035$), which indicates that the effect of pre-play communication decays slightly. Without communication, larger declines of approximately 7% and 16% are observed for TN05 and TN20 respectively ($p_{TN05} = .014$, $p_{TN20} < .001$). Pre-play communication stabilizes cooperation rates over the rounds. There is no decline in the first five rounds for TN35 ($p_{TN35} = .063$). One possible reason is that the cooperation rates are low in TN35, and there is not much room for further decline.

Difference-in-differences tests indicate that the decline in cooperation rates is steeper for TN20 than TN35 ($p_{TN05vs.TN20} = .062$, $p_{TN20vs.TN35} < .001$). However, when communication is possible, the effect of communication does not decay more quickly in treatments with higher noise ($p_{TP05vs.TP20} = .357$, $p_{TP20vs.TP35} = .328$).

Result 1: Pre-play communication increases the average cooperation rates and the initial round cooperation rates at all noise levels. It diminishes the differences in cooperation rates between treatments. The cooperation sustaining effect of communication is similar at different noise levels.

Result 2: Pre-play communication further stabilizes cooperation rates over the rounds. The effect of communication decays slightly over the rounds, and the rate of decay is similar between the treatments.

4.6.2 Conditional Cooperation and Strategy Frequency Estimation

This subsection examines how the cooperation rates differ after different memory-one histories. The histories are action-signal pairs of the previous round, and consist of own action a_i and signal ω_j regarding the opponent's action. There are five possible memory-one states ($\emptyset$, Cc, Cd, Dc, Dd). $\emptyset$ denotes the initial round. The first letter in the other four states represents own action in the previous round, and the second represents the signal regarding the opponent's action. For example, Cc is the case when the subject chose C and received the signal c in the previous round. I denote the probability of cooperation after each history by σ. The corresponding vector of cooperation rates after the five histories is thus ($\sigma_{\emptyset}$, σ_{Cc}, σ_{Cd}, σ_{Dc}, σ_{Dd}). Subjects may also condition their action on the signal regarding their own action, which yields a memory-one history vector with nine states ($\emptyset$, Ccc, Ccd, Cdc, Cdd, Dcc, Dcd, Ddc, Ddd). Results conditional on the nine states are presented in the Appendix D.2.

Table 4.4: Cooperation rates after memory-one histories

	$\sigma_{\emptyset}$	σ_{Cc}	σ_{Cd}	σ_{Dc}	σ_{Dd}	ln$L$
TP05	0.83	0.96	0.64	0.40	0.22	-218.97
	(0.03)	(0.02)	(0.06)	(0.10)	(0.08)	
TP20	0.74	0.92	0.80	0.32	0.24	-287.42
	(0.06)	(0.03)	(0.05)	(0.07)	(0.07)	
TP35	0.66	0.89	0.72	0.28	0.12	-235.83
	(0.06)	(0.03)	(0.05)	(0.07)	(0.04)	
TN05	0.67	0.91	0.62	0.22	0.37	-243.64
	(0.07)	(0.04)	(0.07)	(0.06)	(0.10)	
TN20	0.53	0.78	0.46	0.14	0.21	-267.64
	(0.08)	(0.09)	(0.07)	(0.05)	(0.04)	
TN35	0.27	0.97	0.71	0.10	0.08	-116.73
	(0.09)	(0.05)	(0.09)	(0.05)	(0.03)	

Note: Estimated probabilities of cooperation after memory-one histories. The numbers in the brackets are the standard errors. ln$L$ is the log likelihood of the model.

According to Table 4.4, pre-play communication does not increase cooperation after the mutual cooperation state Cc at $\epsilon = 0.05$ and $\epsilon = 0.35$ ($p_{TP05vs.TN05} = .118$, $p_{TP35vs.TN35} = .102$). At $\epsilon = 0.20$ subjects are more cooperative after Cc with communication ($p_{TP20vs.TN20} = .041$). σ_{Cd} is the probability of cooperation after playing C and receiving a defective signal in the previous round. A higher σ_{Cd} implies that defection does not immediately trigger upon a bad signal, which indicates a higher level of leniency. Overall, leniency increases as noise increases. In treatments without communication, leniency is similar between TN05 and TN20 ($p_{TN05vs.TN20} = .058$), but increases in TN35 compared to TN20 ($p_{TN20vs.TN35} = .018$). With communication, subjects are slightly more lenient in TP20 compared to TP05 ($p_{TP05vs.TP20} = .026$), and they are as lenient in TP20 as in TP35 ($p_{TP20vs.TP35} = .153$). Pre-play communication substantially improves leniency at $\epsilon = 0.20$ ($p_{TP20vs.TN20} < .001$), but not at $\epsilon = 0.05$ and $\epsilon = 0.35$ ($p_{TP05vs.TN05} = .407$, $p_{TP35vs.TN35} = .428$). σ_{Dc} and σ_{Dd} measure whether subjects are willing to return to cooperation. Higher σ_{Dc} and σ_{Dd} indicate higher levels of forgiveness. Without pre-play communication, the level of forgiveness is lower at $\epsilon = 0.35$ ($p_{TN20vs.TN35} = .031$, $p_{TN05vs.TN35} = .002$). The greater leniency and lower forgiveness in TN35 is puzzling. With pre-play communication, forgiveness level is similar at the three noise levels ($p_{TP05vs.TP20} = .405$,$p_{TP20vs.TP35} = .139$, $p_{TP05vs.TP35} = .120$). Pre-play communication increases forgiveness level at $\epsilon = 0.35$ ($p_{TP35vsTN35} = .024$), but not at $\epsilon = 0.05$ and $\epsilon = 0.20$ ($p_{TP05vsTN05} = .433$, $p_{TP20vsTN20} = .077$).

Table 4.5: Frequencies of pure strategies

	TP05	TP20	TP35	TN05	TN20	TN35
ALLC	0.24	0.40	-	0.37	-	0.12
	(0.11)	(0.11)		(0.12)		(0.08)
ALLD	-	0.20	0.20	0.25	0.51	0.68
		(0.07)	(0.10)	(0.09)	(0.11)	(0.12)
DC	0.03	-	-	0.04	-	-
	(0.03)			(0.04)		
GRIM	0.22	-	0.16	-	0.12	-
	(0.10)		(0.12)		(0.09)	
GRIM2	0.45	0.37	0.31	0.21	0.14	-
	(0.13)	(0.11)	(0.15)	(0.10)	(0.08)	
GRIM3	-	-	0.17	0.14	0.16	0.06
			(0.12)	(0.11)	(0.09)	(0.06)
DGRIM2	-	-	-	-	-	0.07
						(0.07)
WSLS	0.06		-	0.08	-	-
	(0.06)			(0.07)		
TF2T	-	-	0.16	-	-	0.06
			(0.10)			(0.06)
DTF3T	-	0.03	-	-	-	-
		(0.03)				
γ	0.11	0.12	0.19	0.13	0.18	0.08
BIC	500.02	544.13	528.63	456.09	512.30	226.98
$\ln L$	-241.35	-265.13	-256.37	-218.5	-249.79	-106.56

Note: This table reports the maximum-likelihood shares of pure strategies. The estimation procedure assumes constant strategy use across all supergames. γ is the estimated tremble probability, which avoids likelihood shares of zero when subjects deviate from a choice pattern. Strategies are selected based on Bayesian Information Criterion. Strategies attracting zero shares are omiited (-). Standard errors are reported in parentheses. Values may not add up to 1 because of rounding.

I use the strategy frequency estimation method based on Dal Bó and Fréchette (2011) to further analyze which strategies subjects play out of the 20 pure candidate strategies from Fudenberg et al. (2012). The results are presented in Table 4.5. With pre-play communication, subjects do not play ALLD (always defect) at noise $\epsilon = 0.05$. Without communication, the shares of ALLD increase as noise increases. The shares of ALLC (always cooperate) is high in TP20, but ALLC is not played in TN20. GRIM is a grim-trigger strategy which triggers defection immediately after observing a bad signal. GRIM2 and GRIM3 are the lenient versions of GRIM. Defection is triggered after two or three consecutive bad signals. The shares of GRIM2 and GRIM3 are higher in TP35 compared to TP20. Subjects in TP35 also play the lenient and forgiving strategy TF2T (tit-for-2-tats). This is a tit-for-tat strategy but defection will not be played unless the signal is defective in both of the last two rounds. Without communication, subjects play less lenient strategies as noise increases. These estimation results indicate that, with pre-play communication, subjects are less likely to immediately enter the punishment state upon a defective signal, and many subjects do not punish. This is not observed in treatments without communication. Other strategies such as DC (alternator, start with D), DGRIM2 (D-to-GRIM2), WSLS (win-stay-lose-shift) and DTF3T (D-to-tit-for-3-tats) are also played with low frequencies. Note that in Table 4.4, compared with TN05 and TN20, leniency is greater in TN35, but forgiveness is lower. This may be due to the higher proportion of subjects playing ALLC and ALLD in TN35. However, it is still surprising that ALLC accounts for 12% towards the strategy shares in TN35. One of the key questions is what beliefs subjects hold about their opponent's strategies. I examine their beliefs in the next subsection.

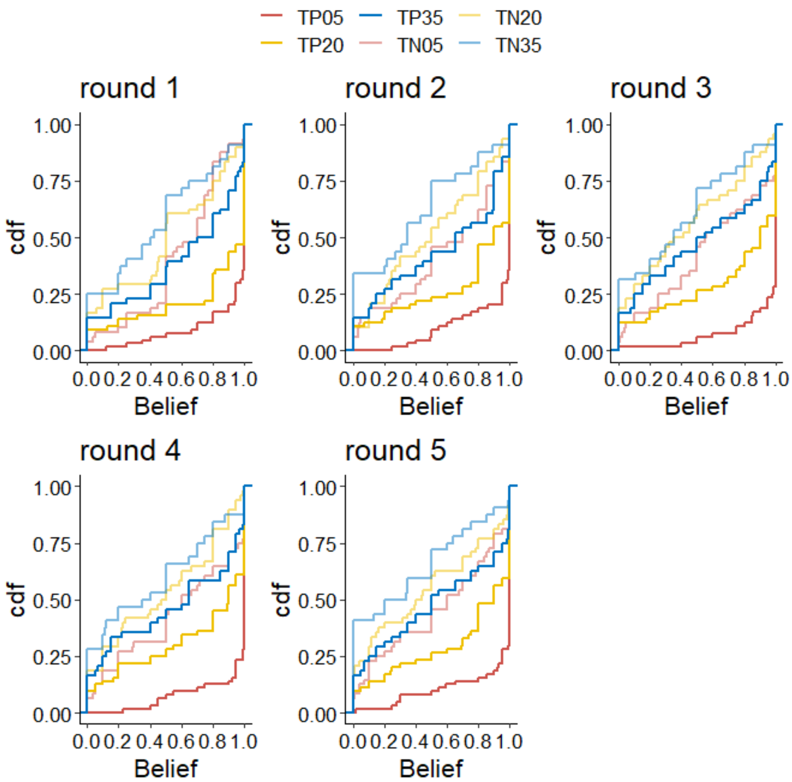
Result 3: Leniency increases as noise increases. Communication increases leniency and forgiveness in strategy choices.

4.6.3 Beliefs

I start by examining the cumulative distribution of beliefs in the first five rounds of supergames five and six. As Figure 4.4 shows, the evolution of beliefs is similar over the rounds. Beliefs about cooperation are higher in treatments with lower noise levels. This result indicates that pre-play communication reduces the strategic uncertainty towards cooperating by making subjects more optimistic that their opponents will cooperate with a higher probability. This effect is not short-lived. Subjects in the pre-play communication treatments still have highers belief about cooperation in the subsequent rounds.

I take a closer look at the round one belief because beliefs in the initial round are not affected by the history of play. Subjects form their beliefs in the first round either by coordinating with their opponents, or through introspection when communication is impossible. Round one belief shows the initial assessment of strategic uncertainty towards cooperation. With pre-play communication, subjects' beliefs are higher when the noise level is lower ($p_{TP05vs.TP20} = .032$,

Figure 4.4: Belief distribution in the first five rounds

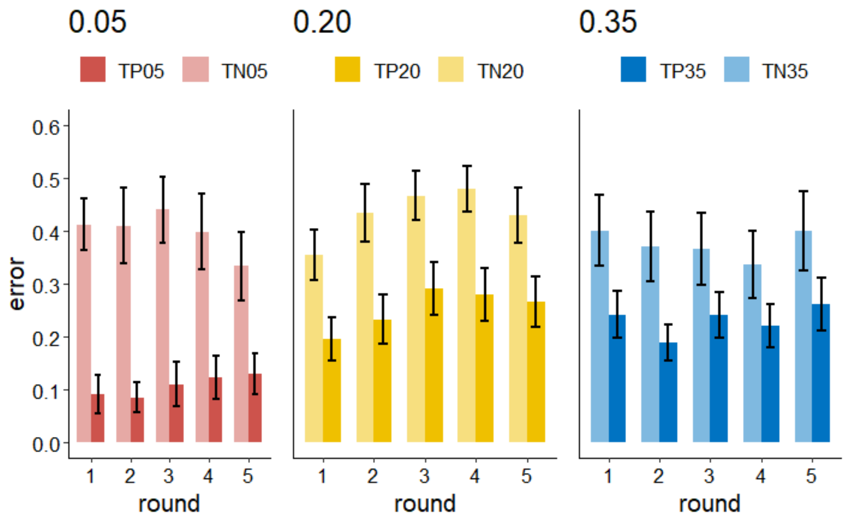


Note: Cumulative distribution of beliefs in the first five rounds of supergames five and six.

$p_{TP20vs.TP35} = .031$, $p_{TP05vs.TP35} < .001$). The average initial belief in TP05 is 12% higher than that in TP20 and 29% higher than that in TP35. Without pre-play communication, subjects report similar beliefs between TN05 and TN20 ($p_{TN05vs.TN20} = .151$) and between TN20 and TN35 ($p_{TN20vs.TN35} = .232$). The average belief in TN05 is 18% higher than that in TN35 ($p_{TN20vs.TN35} = .044$). The reported initial beliefs are around 20% to 30% higher in treatments with pre-play communication at every noise level ($p_{TP05vs.TN05} < .001$, $p_{TP20vs.TN20} = .001$, $p_{TP35vs.TN35} = .040$). The difference-in-differences tests show that the belief-increasing effect from communication does not differ between the different noise levels ($p_{0.05vs.0.20} = .408$, $p_{0.20vs.0.35} = .277$, $p_{0.05vs.0.35} = .195$).

One interesting question is whether subjects can correctly predict the actions of their opponents. I expect that it is easier to make correct predictions when subjects pre-play communicate, because they can coordinate with their opponents. However, this effect may decay because the average cooperation rate declines over the rounds. Moreover, subjects do not receive feedback on their opponent's actual choice, which makes it even more difficult to infer correctly from past signals what actions their opponent will take in the current round.

Figure 4.5: Error in belief



Note: This figure shows the average absolute error in belief at every noise level. The error bars are the standard errors which are two-way clustered on subject level and match level.

According to Figure 4.5, errors occur in all supergames, but are higher in supergames without communication. The error rates are similar at the three noise levels in the no-communication treatments and are similar between TP30 and TP35. Subjects make more accurate predictions in treatment TP05. They do not correct their predictions over the rounds. Their predictions are not more erroneous in the later rounds.

Another interesting question is whether subjects adjust their beliefs according to the signals. Are they more likely to believe that their opponent will cooperate after a cooperative signal? And does the adjustment differ according to the round one histories? I examine this question in the Appendix D.4. For every treatment, I look at the conditional beliefs after four round-one histories. The histories are action-signal pairs. I assume that the subjects condition their beliefs on their action and the signal representing the opponent's action. According to Figures D.4.1 to D.4.6, subjects do not their beliefs to signals. The evolution of their beliefs is stabler than the evolution of the signals. Pre-play communication helps to raise the subjective beliefs in almost all cases at any noise level. Beliefs are higher when the subject's own choice in round one is C. This observation is not unique to treatments with pre-play communication. While subjects in the pre-play communication treatments can coordinate strategies such as always-cooperation or always-defection, subjects do not have the opportunity to coordinate their strategies if they do not engage in pre-play communication. However, subjects in the no communication treatments still hold high beliefs in the case Cd in which they themselves cooperate in round one but receive a defective signal. One possible explanation is that they tend to overestimate the probability of other players choosing the same action. This may also explain why the estimated proportion of ALLC is not low in TN35. According to Figure D.4.6, subjects have high beliefs following history Cd; however, they overestimate the likelihood that their opponent plays the same action.

Result 4: The subjective beliefs about cooperation are higher with communication and are higher when the noise level is low. Beliefs are less accurate in treatments without communication.

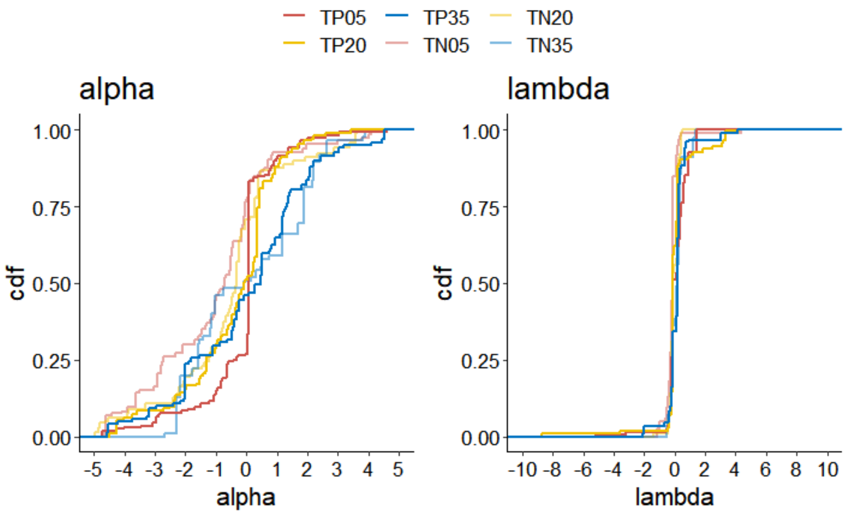
Result 5: Beliefs remain stable over the rounds and they are history-dependent. Subjects hold higher beliefs if their own choice in round one is to cooperate.

This study makes a distinction between beliefs and attitudes towards strategic uncertainty. In the next subsection, I estimate uncertainty attitudes.

4.6.4 Attitudes Towards Strategic Uncertainty

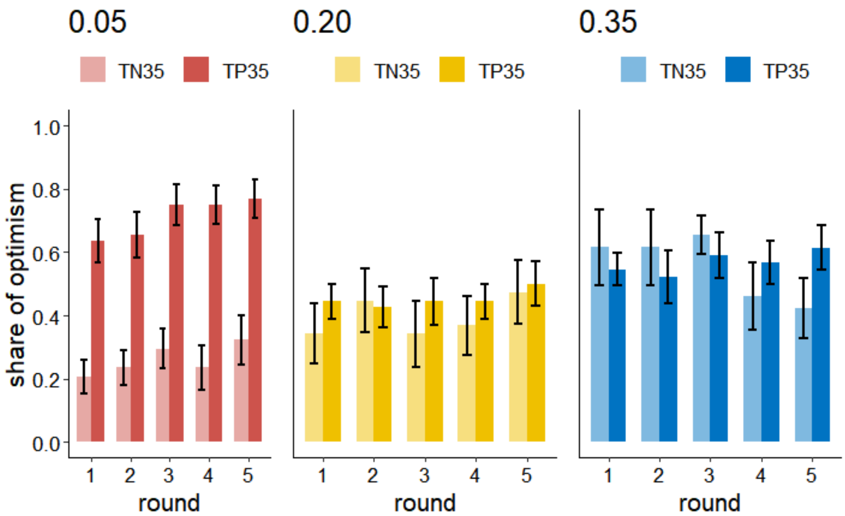
The estimation of attitudes towards strategic uncertainty follows a two-step procedure. First, I estimate the individual relative risk aversion parameter r from the self-reported certainty equivalents of the 22 lotteries. This estimation method follows Hey et al. (2009), in which the authors calculate the difference between the stated certainty equivalence and the true certainty equivalence based on the expected utility of the lottery. This difference is assumed to follow a normal distribution with a mean of zero and standard deviation s. The relative risk aversion parameter r and the standard deviation s are obtained from the maximum likelihood. According to Charness et al. (2019), the range of relative risk aversion is found between -3 and 3. The sample is restricted to subjects whose estimated r is in [-3,3]. The full sample comprises 152 subjects, and the restricted sample size is 124. In the second step, I plug the estimated r into equations 4.9 and 4.7 to calculate (α, λ) for each subject in each round.

Figure 4.6: Distribution of α and λ



Note: This figure shows the empirical cumulative distributions of α and $lambda$. α is trimmed to $[-5, 5]$ because of outliers. $lambda$ on the right graph is neglog-transformed $sign(\lambda)/log(1 + |\lambda|)$ because of the wide range, and the values are trimmed to $[-10, 10]$.

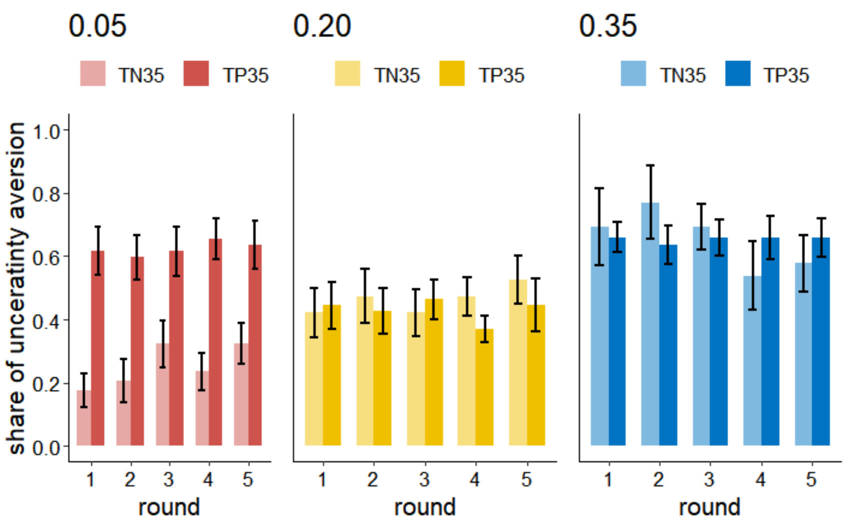
Figure 4.7: Proportions of rounds in which subjects are optimistic



Note: This figure shows the proportion of cases in which subjects are optimistic about their beliefs. The samples are restricted on the relative risk aversion parameter r.

The median of the estimated relative risk aversion parameter is -0.44, with 37 risk-averse

Figure 4.8: Proportions of rounds in which subjects are uncertainty averse



Note: This figure shows the proportion of cases in which subjects are uncertainty averse. The samples are restricted based on the relative risk aversion parameter r.

subjects and 87 risk-loving subjects.[7] The sign test shows that there are more risk-loving subjects ($p < .001$). The medians of the estimated α and λ are -0.09 and 0, respectively. α captures optimism in the desired outcome. A positive α means that subjects are optimistic that to receive the desired expected outcome. λ is the uncertainty aversion parameter that captures the degree of aversion against strategic uncertainty. A positive λ means that the subject has dis-utility from the source of uncertainty. The estimated α is positive in 599 of 1240 rounds, and the estimated λ is positive in 621 of 1240 rounds. The sign tests show that there are not more cases in which subjects are pessimistic ($p = .244$) and not more cases in which they have aversion against uncertainty ($p = .689$).

Figure 4.6 shows the cumulative distributions of α and λ based on the restricted r. The ranges of the two estimated parameters are wide.[8] In the figure, α is trimmed to $[-5, 5]$ and λ is neglog-

[7]It is possible that some subjects did not fully understand the BDM mechanism used for the elicitation tasks. In fact, a number of subjects reported finding the elicitation tasks difficult to comprehend at the end of the sessions. As a result, the estimates of attitudes towards strategic uncertainty obtained in this subsection may not accurately reflect the true attitudes of all subjects.

[8]The wide range is due to the parameterization of Prisoners' Dilemma. The difference in the certainty equivalents regarding lottery C and lottery D should be explained by the difference in the expected utility if the subject views the two lotteries in a Bayesian way. If the subject does not adopt a Bayesian view, there is unexplained difference. According to the theoretical model, this unexplained difference can be explained by the difference in $(EU(C) - EU(D))$. α is the weight that the subject assigns to obtaining the expected utility

transformed and trimmed to $[-10, 10]$. According to Figure 4.6, subjects are less pessimistic in treatment TP05. Subjects in TN05 are more pessimistic than those in the other treatments. I then focus on the proportion of cases in which subjects are optimistic and the proportion of cases in which subjects are uncertainty averse. The results are shown in Figures 4.7 and 4.8. Both figures are based on the restricted r. The unrestricted version is presented in the Appendix D.5.

Compared with TN05, there are 49% more cases in treatment TP05 in which subjects are optimistic about their beliefs ($p_{TP05vs.TN05} < .001$). No difference is found between TP20 and TN20 ($p_{TP20vs.TN20} = .293$), neither is difference found between TP35 and TN35 ($p_{TP35vs.TN35} = .440$). This implies that subjects are more optimistic about obtaining the higher expected utility of partner playing cooperation. However, this effect is restricted to the low noise level. For treatments with pre-play communication, there are 25% more cases in TP05 in which subjects are more optimistic about receiving the higher expected outcome than in TP20 ($p_{TP05vs.TP20} = .015$). The proportion is similar between TP20 and TP35 ($p_{TP20vs.TP35} = .095$), and similar between TP05 and TP35 ($p_{TP05vs.TP35} = .224$). Subjects are the most optimistic when the noise is at the lowest level when communication is possible. In treatments without pre-play communication, although the mean of the reported beliefs is the lowest in TN35, there is a higher proportion of cases in which subjects are more optimistic that their opponent will cooperate in TN35 than in TN05 ($p_{TN05vs.TN35} = .003$). It is unclear why optimism increases when the noise is high. The proportion of optimism is similar between TN05 and TN20 ($p_{TN05vs.TN20} = .073$) and similar between TN20 and TN35 ($p_{TN20vs.TN35} = .080$).

As for the share of positive λ, pre-play communication increases the share of cases in which subjects are uncertainty averse at the noise level $\epsilon = 0.05$ ($p_{TP05vs.TN05} < .001$, $p_{TP20vs.TN20} = .436$, $p_{TP35vs.TN35} = .500$). This finding is surprising and it is unclear why subjects have higher dis-utility from the source of uncertainty when communication is possible. Pre-play communication reduces the proportion of uncertainty aversion in TP20 compared to TP35 by approximately 25% ($p_{TP20vs.TP35} = .019$). The proportions are similar between TP05 and TP20 ($p_{TP05vs.TP20} = .067$) and similar between TP05 and TP35 ($p_{TP05vs.TP35} = .253$). In treatments without communication, the proportion of cases in which subjects are uncertainty averse is the

when the opponent cooperates rather than defects. $(EU(C) - EU(D))$ is small due to the nature of Prisoners' Dilemma. The small difference in $(EU(C) - EU(D))$ makes α sensitive to the size of the unexplained difference.

highest in TN35. It is 50% higher than in TN05 ($p_{TN05vs.TN35} < .001$), and 23% higher than in TN20 ($p_{TN20vs.TN35} = .056$).

I further check the effect of the interaction of communication and noise on whether subjects are optimistic or uncertainty averse. Pre-play communication substantially increases the proportion of optimism at $\epsilon = 0.05$ compared to $\epsilon = 0.20$ ($p_{0.05vs.0.20} = .006$) and $\epsilon = 0.35$ ($p_{0.05vs.0.35} = .005$). Similar results are observed for uncertainty aversion. Pre-play communication increases the proportion of cases in which subjects are uncertainty averse at $\epsilon = 0.05$ compared to $\epsilon = 0.20$ ($p_{0.05vs.0.20} = .002$), and it reduces the proportion of uncertainty aversion at $\epsilon = 0.05$ compared to $\epsilon = 0.35$ ($p_{0.05vs.0.35} = .007$).

In the Appendix D.6, I examine the changes in the shares of positive α and λ conditional on the histories in the initial round. The changes in subjects' attitudes towards strategic uncertainty do not seem to relate to the changes in beliefs and signals. Subjects are neither more optimistic nor less uncertainty averse when they receive a cooperative signal in the initial round.

Result 6: Pre-play communication makes subjects more optimistic that their opponent will cooperate at the low noise level. Pre-play communication also makes the subjects more likely to be averse against uncertainty when the noise level is low. Changes in attitudes towards uncertainty are unrelated to the history of the initial round.

4.7 Conclusion

Indefinitely repeated economic interactions with imperfect monitoring structures pose a challenge to sustaining cooperation. Past literature suggests that pre-play communication can be a key determinant of cooperation in such scenarios. One possible mechanism through which communication affects cooperation is by reducing perceived strategic uncertainty. This study presents empirical evidence on the cooperation-sustaining effect of pre-play communication in scenarios where cooperation is risky or cooperative equilibrium cannot be constructed. I conduct laboratory experiments using the indefinitely repeated Prisoners' Dilemma game with imperfect public monitoring. Subjective beliefs are elicited to measure perceived strategic uncertainty towards cooperation, while attitudes towards strategic uncertainty are also estimated. This study uses a structural model that allows for the distinguish between the perceived strategic uncertainty and uncertainty attitudes.

I conduct six treatments that vary in two dimensions: the noise levels concerning how accurately the subjects monitor each other's actions and communication opportunities. The results show that, despite the high risk of cooperation in high noise levels, the average cooperation rate is still high if subjects engage in a pre-play communication stage. Furthermore, the effect of pre-play communication is not limited to scenarios where cooperative equilibrium can be established but extends to scenarios where cooperation is risk or cooperative equilibrium cannot be constructed. Communication increases leniency and forgiveness of the chosen strategies.

One potential explanation for how pre-play communication affects cooperation and strategy choice is by reducing strategic uncertainty, which refers to the uncertainty that a subject has about the likelihood of their opponent cooperating. To measure subjective beliefs about cooperation, I elicit these beliefs in all treatments and find that they are low when the noise level is high. However, when pre-play communication is allowed, the elicited beliefs are higher. Additionally, the round-one beliefs with pre-play communication are higher than those without, indicating that subjects coordinate on playing cooperation via communication. These findings suggest that pre-play communication reduces strategic uncertainty by increasing subjective beliefs about cooperation.

Another possible channel through which communication may come into play is by influencing subjects' attitudes towards strategic uncertainty. This attitude has two dimensions. The first dimension is an extra optimism or pessimism that the subject associates to the desirable expected outcome. The second dimension is the preference for, or aversion against uncertainty. Certainty equivalents for action C and action D are elicited in each round, and the attitudes towards strategic uncertainty are estimated for each subject through a structural model. Subjects in the treatment with pre-play communication of low noise level are found to be more optimistic that their partner will cooperate, comparing to subjects in the treatment of the same noise level but without communication. This observation is limited to the low noise level though. The findings raise several questions. One puzzling result is the higher proportion of subjects who are averse to uncertainty when they are allowed to communicate before the game. It is unclear why this should be the case. Additionally, it is unclear why the availability of pre-play communication increases subjects' optimism about their partner's cooperation when noise levels are high. It is possible that some subjects did not fully understand the elicitation mechanism used in this study, which may have impacted the accuracy of the estimation results.

To obtain a more accurate understanding of subjects' attitudes towards uncertainty, future research could consider using simpler elicitation designs.

Despite the deviations observed, this paper provides confirmation of the crucial role that communication plays. The effect of pre-play communication extends beyond the scenarios where cooperative equilibrium can be constructed, and it also increases the leniency and forgiveness of chosen strategies. Subjective beliefs are shown to be affected by both the noise level and communication opportunity. In particular, communication reduces strategic uncertainty by increasing subjective beliefs about cooperation.

Possible avenues for future research could include extending the framework to private monitoring scenarios, or exploring the effect of different communication structures, such as repeated communication, on perceived strategic uncertainty. Such investigations may further deepen our understanding of how communication can influence cooperation in economic interactions.

Bibliography

Adamopoulos, P. and Tuzhilin, A. (2014). On unexpectedness in recommender systems: Or how to better expect the unexpected. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(4):1–32.

Addicott, M. A., Pearson, J. M., Sweitzer, M. M., Barack, D. L., and Platt, M. L. (2017). A primer on foraging and the explore/exploit trade-off for psychiatry research. *Neuropsychopharmacology*, 42(10):1931–1939.

Adomavicius, G., Huang, Z., and Tuzhilin, A. (2008). Personalization and recommender systems. In *State-of-the-Art Decision-Making Tools in the Information-Intensive Age*, pages 55–107. INFORMS.

Adomavicius, G. and Tuzhilin, A. (2005a). Personalization technologies: a process-oriented perspective. *Communications of the ACM*, 48(10):83–90.

Adomavicius, G. and Tuzhilin, A. (2005b). Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17:734–749.

Aoyagi, M., Bhaskar, V., and Fréchette, G. R. (2019). The impact of monitoring in infinitely repeated games: Perfect, public, and private. *American Economic Journal: Microeconomics*, 11(1):1–43.

Aoyagi, M. and Fréchette, G. R. (2009). Collusion as public monitoring becomes noisy: Experimental evidence. *Journal of Economic Theory*, 144(3):1135–1165.

Aoyagi, M., Frechette, G. R., and Yuksel, S. (2022). Beliefs in repeated games. *ISER DP No. 1119RR, Available at SSRN: https://ssrn.com/abstract=3786422 or http://dx.doi.org/10.2139/ssrn.3786422*.

Awaya, Y. and Krishna, V. (2016). On Communication and Collusion. *American Economic Review*, 106(2):285–315.

Awaya, Y. and Krishna, V. (2019). Communication and cooperation in repeated games. *Theoretical Economics*, 14(2):513–553.

Backhaus, T. and Breitmoser, Y. (2018). God does not play dice, but do we? Technical Report 190, Rationality and Competition Discussion Paper Series 96.

Baillon, A., Huang, Z., Selim, A., and Wakker, P. P. (2018). Measuring ambiguity attitudes for all (natural) events. *Econometrica*, 86(5):1839–1858.

Bandeen-Roche, K., Miglioretti, D. L., Zeger, S. L., and Rathouz, P. J. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association*, 92(440):1375–1386.

Barro, R. and Gordon, D. B. (1983). Rules, discretion, and reputation in a model of monetary policy. *Journal of Monetary Economics*, 12(1):101–121.

Becker, G. M., Degroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3):226–232.

Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1):122–142.

Biernacki, C., Celeux, G., and Govaert, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:719–725.

Blonski, M., Ockenfels, P., and Spagnolo, G. (2011). Equilibrium Selection in the Repeated Prisoner's Dilemma: Axiomatic Approach and Experimental Evidence. *American Economic Journal: Microeconomics*, 3(3):164–192.

Bock, O., Baetge, I., and Nicklisch, A. (2014). hroot: Hamburg registration and organization online tool. *European Economic Review*, 71:117 – 120.

Breitmoser, Y. (2015). Cooperation, but no reciprocity: Individual strategies in the repeated prisoner's dilemma. *American Economic Review*, 105(9):2882–2910.

Brown, V. M., Hallquist, M. N., Frank, M. J., and Dombrovski, A. Y. (2022). Humans adaptively resolve the explore-exploit dilemma under cognitive constraints: Evidence from a multi-armed bandit task. *Cognition*, 229:105233.

Bruttel, L., Bulutay, M., Cornand, C., Heinemann, F., and Zylbersztejn, A. (2023). Measuring strategic uncertainty attitudes (forthcoming). *Experimental Economics*.

Bursian, D. and Fürth, S. (2015). Trust me! i am a european central banker. *Journal of Money, Credit and Banking*, 47(8):1503–1530.

Camerer, C. and Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, 5(4):325–370.

Cameron, a. C., Gelbach, J. B., and Miller, D. L. (2011). Robust Inference With Multiway Clustering. *Journal of Business & Economic Statistics*, 29(2):238–249.

Capponi, A., Olafsson, S., and Zariphopoulou, T. (2022). Personalized robo-advising: Enhancing investment through client interaction. *Management Science*, 68(4):2485–2512.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, 76(1).

Castells, P., Hurley, N., and Vargas, S. (2022). Novelty and diversity in recommender systems. In *Recommender systems handbook*, pages 603–646. Springer.

Chaharsooghi, S. K., Heydari, J., and Zegordi, S. H. (2008). A reinforcement learning model for supply chain ordering management: An application to the beer game. *Decision Support Systems*, 45(4):949–959.

Chakroun, K., Mathar, D., Wiehler, A., Ganzer, F., and Peters, J. (2020). Dopaminergic modulation of the exploration/exploitation trade-off in human decision-making. *Elife*, 9:e51260.

Charness, G., Garcia, T., Offerman, T., and Villeval, M. C. (2019). Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory? Working Papers 1921, Groupe d'Analyse et de Théorie Economique Lyon St-Étienne (GATE Lyon St-Étienne), Université de Lyon.

Compte, O. (1998). Communication in Repeated Games with Imperfect Private Monitoring. *Econometrica*, 66(3):587–626.

Compte, O. and Postlewaite, A. (2015). Plausible cooperation. *Games and Economic Behavior*, 91:45–59.

Cooper, D. J. and Kühn, K.-U. (2014). Communication, renegotiation, and the scope for collusion. *American Economic Journal: Microeconomics*, 6(2):247–278.

Cooper, R. W., DeJong, D. V., Forsythe, R., and Ross, T. W. (1992). Communication in Coordination Games. *Quarterly Journal of Economics*, 107(2):739–771.

Cremer, J. (1986). Cooperation in ongoing organizations. *Quarterly Journal of Economics*, 101(1):33–49.

Dal Bo, E. (2006). Committees with supermajority voting yield commitment with flexibility. *Journal of Public Economics*, 90(4):573–599.

Dal Bó, P. and Fréchette, G. R. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, 101(1):411–429.

Dal Bó, P. and Fréchette, G. R. (2018). On the determinants of cooperation in infinitely repeated games: A survey. *Journal of Economic Literature*, 56(1):60–114.

Danwitz, L., Mathar, D., Smith, E., Tuzsus, D., and Peters, J. (2022). Parameter and model recovery of reinforcement learning models for restless bandit problems. *Computational Brain & Behavior*, pages 1–17.

Danz, D., Vesterlund, L., and Wilson, A. J. (2022). Belief elicitation and behavioral incentive compatibility. *American Economic Review*, 112(9):2851–83.

Daw, N. D., O'doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095):876–879.

Dayton, C. M. and Macready, G. B. (1988). Concomitant-variable latent-class models. *Journal of the American Statistical Association*, 83(401):173–178.

Dempster, A., Laird, N., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B*, 39(1):1–38.

Dietvorst, B. J. and Bharti, S. (2020). People reject algorithms in uncertain decision domains because they have diminishing sensitivity to forecasting error. *Psychological science*, 31(10):1302–1314.

Dietvorst, B. J., Simmons, J. P., and Massey, C. (2015). Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1):114.

Durlauf, S. and Fafchamps, M. (2004). Social capital. NBER Working Paper No. 10485.

Dvorak, F. (2018). stratest: An r package for strategy estimation in experimental economics. Technical report, SSRN working paper.

Dvorak, F. (2019). stratest: Strategy estimation in r. Manuscript.

Dvorak, F. and Fehrler, S. (2018). Negotiating cooperation under uncertainty. *IZA working paper*.

Dvorak, F. and Fehrler, S. (2019). Negotiating Cooperation under Uncertainty: Communication in Noisy, Indefinitely Repeated Interactions. *Mimeo*.

Ecken, P. and Pibernik, R. (2016). Hit or miss: What leads experts to take advice for long-term judgments? *Management Science*, 62(7):2002–2021.

Ehrmann, M., Soudan, M., and Stracca, L. (2013). Explaining european union citizens' trust in the european central bank in normal and crisis times. *Scandinavian Journal of Economics*, 115(3):781–807.

Ellingsen, T. and Östling, R. (2010). When Does Communication Improve Coordination? *American Economic Review*, 100(4):1695–1724.

Embrey, M., Fréchette, G., and Stacchetti, E. (2013). An Experimental Study of Imperfect Public Monitoring: Efficiency Versus Renegotiation-Proofness. *Mimeo*.

Embrey, M., Frechette, G. R., and Yuksel, S. (2017). Cooperation in the Finitely Repeated Prisoner 's Dilemma. *Quaterly Journal of Economics*, (forthc).

Engle-Warnick, J. and Slonim, R. L. (2004). The evolution of strategies in a repeated trust game. *Journal of Economic Behavior & Organization*, 55(4):553–573.

Engle-Warnick, J. and Slonim, R. L. (2006). Inferring repeated-game strategies from actions: Evidence from trust game experiments. *Economic Theory*, 28(3):603–632.

Enke, B. and Zimmermann, F. (2017). Correlation Neglect in Belief Formation. *The Review of Economic Studies*, 86(1):313–332.

Falk, A. and Zimmermann, F. (2017). Consistency as a signal of skills. *Management Science*, 63(7):2197–2210.

Farvaque, Etienne, M. A. H. and Mihailov, A. (2012). Who supports the ecb? evidence from eurobarometer survey data. Manuscript.

Fehrler, S. and Hughes, N. (2018). How transparency kills information aggregation: Theory and experiment. *American Economic Journal: Microeconomics*, 10.

Filiz, I., Judek, J. R., Lorenz, M., and Spiwoks, M. (2021). Reducing algorithm aversion through experience. *Journal of Behavioral and Experimental Finance*, 31:100524.

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

Fischer, J. and Hahn, V. (2008). Determinants of trust in the european central bank. Working Paper Series in Economics and Finance 695, Stockholm School of Eco- nomics.

Fréchette, G. R. and Yuksel, S. (2017). Infinitely repeated games in the laboratory: four perspectives on discounting and random termination. *Experimental Economics*, 20(2):279–308.

Fudenberg, D., Rand, D. G., and Dreber, A. (2012). Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World. *American Economic Review*, 102(2):720–749.

Ganju, K. K., Atasoy, H., McCullough, J., and Greenwood, B. (2020). The role of decision support systems in attenuating racial biases in healthcare delivery. *Management science*, 66(11):5171–5181.

Gershman, S. J. (2020). Origin of perseveration in the trade-off between reward and complexity. *Cognition*, 204:104394.

Gershman, S. J., Horvitz, E. J., and Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245):273–278.

Gino, F., Shang, J., and Croson, R. (2009). The impact of information from similar or different advisors on judgment. *Organizational behavior and human decision processes*, 108(2):287–302.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., and Soutter, C. L. (2000). Measuring trust. *The Quarterly Journal of Economics*, 115(3):811–846.

Green, E. J. and Porter, R. H. (1984). Noncooperative Collusion under Imperfect Price Information. *Econometrica*, 52(1):87.

Greiner, B. (2004). The online recruitment system orsee - a guide for the organization of experiments in economics. Papers on strategic interaction, Max Planck Institute of Economics, Strategic Interaction Group.

Greiner, B. (2023). Strategic uncertainty aversion in bargaining — experimental evidence. *Journal of Economic Psychology*, 95:102604.

Gros, D. and Roth, F. (2010). The financial crisis and citizens' trust in the european central bank. CEPS Working Document No. 334.

Guiso, L., Sapienza, P., and Zingales, L. (2004). The role of social capital in financial development. *American Economic Review*, 94(3):526–556.

Gächter, S. and Falk, A. (2002). Reputation and reciprocity: Consequences for the labour relation. *The Scandinavian Journal of Economics*, 104(1):1–26.

Hansen, S. and McMahon, M. (2016). First impressions matter: Signalling as a source of policy dynamics. *The Review of Economic Studies*, 83(4):1645–1672.

Harsanyi, J. C. and Selten, R. (1988). *A General Theory of Equilibrium Selection in Games*. MIT Press, Cambridge, MA.

Heinemann, F., Nagel, R., and Ockenfels, P. (2009). Measuring strategic uncertainty in coordination games. *The Review of Economic Studies*, 76(1):181–221.

Heller, Y. (2017). Instability of belief-free equilibria. *Journal of Economic Theory*, 168:261–286.

Hey, J. D., Morone, A., and Schmidt, U. (2009). Noise and bias in eliciting preferences. *Journal of Risk and Uncertainty*, 39(3):213–235.

Himmelstein, M. (2022). Decline, adopt or compromise? a dual hurdle model for advice utilization. *Journal of Mathematical Psychology*, 110:102695.

Hou, Y. T.-Y. and Jung, M. F. (2021). Who is the expert? reconciling algorithm aversion and algorithm appreciation in ai-supported decision making. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2):1–25.

Ihssen, N., Mussweiler, T., and Linden, D. E. (2016). Observing others stay or switch–how social prediction errors are integrated into reward reversal learning. *Cognition*, 153:19–32.

Jussupow, E., Benbasat, I., and Heinzl, A. (2020). Why are we averse towards algorithms? a comprehensive literature review on algorithm aversion.

Kandori, M. and Matsushima, H. (1998). Private Observation, Communication and Collusion. *Econometrica*, 66(3):627–652.

Kartal, M. and Müller, W. (2018). A new approach to the analysis of cooperation under the shadow of the future: Theory and experimental evidence. *SSRN Working Paper*.

Kawaguchi, K. (2021). When will workers follow an algorithm? a field experiment with a retail business. *Management Science*, 67(3):1670–1695.

Kelsey, D. and le Roux, S. (2014). An experimental study on the effect of ambiguity in a coordination game. *Theory and Decision*, 79:667–688.

Knack, S. and Keefer, P. (1997). Does social capital have an economic payoff? a cross-country investigation. *The Quarterly journal of economics*, 112(4):1251–1288.

Kroszner, R. S. and Stratmann, T. S. (2000). Congressional committees as reputation-building mechanisms. *Business and Politics*, 2(1):35–52.

Kugler, T., Bornstein, G., Kocher, M. G., and Sutter, M. (2007). Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy. *Journal of Economic Psychology*, 28(6):646–657.

Kydland, F. E. and Prescott, E. C. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, 85(3):473–492.

La Porta, R., Lopez-de Silanes, F., Shleifer, A., and Vishny, R. W. (1997). Trust in large organizations. *The American Economic Review, Papers& Proceedings*, 87(2):333–338.

Laureiro-Martinez, D., Brusoni, S., Tata, A., and Zollo, M. (2019). The manager's notepad: Working memory, exploration, and performance. *Journal of Management Studies*, 56(8):1655–1682.

Levy, G. and Razin, R. (2015). Correlation neglect, voting behavior, and information aggregation. *American Economic Review*, 105(4):1634–45.

Logg, J. M., Minson, J. A., and Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151:90–103.

Madhavan, P. and Wiegmann, D. A. (2007). Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science*, 8(4):277–301.

Mahmud, H., Islam, A. N., Ahmed, S. I., and Smolander, K. (2022). What influences algorithmic decision-making? a systematic literature review on algorithm aversion. *Technological Forecasting and Social Change*, 175:121390.

Meder, B., Wu, C. M., Schulz, E., and Ruggeri, A. (2021). Development of directed and random exploration in children. *Developmental science*, 24(4):e13095.

Minson, J. A., Liberman, V., and Ross, L. (2011). Two to tango: Effects of collaboration and disagreement on dyadic judgment. *Personality and Social Psychology Bulletin*, 37(10):1325–1338.

Obara, I. (2009). Folk Theorem with Communication. *Journal of Economic Theory*, 144(1):120–134.

Offerman, T., Potters, J., and Verbon, H. A. (2001). Cooperation in an overlapping generations experiment. *Games and Economic Behavior*, 36(2):264–275.

Orth, D. A., Buchanan, M., Amresh, A., Smith, C., Lematta, G., Cooke, N., Fouse, A., and Dubrow, S. (2021). Designing for exploration and exploitation in experimental search and rescue scenarios. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 65, pages 720–725. SAGE Publications Sage CA: Los Angeles, CA.

Panniello, U., Gorgoglione, M., and Tuzhilin, A. (2016). Research note—in carss we trust: How context-aware recommendations affect customers' trust and other business performance measures of recommender systems. *Information Systems Research*, 27:182–196.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., and Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1):195–203.

Pouya, P. and Madni, A. (2021). Performing active search to locate indication of ancient water on mars: An online, probabilistic approach. In *ASCEND 2021*, page 4024.

Prahl, A. and Van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting*, 36(6):691–702.

Pálfi, B., Arora, K., and Kostopoulou, O. (2022). Algorithm-based advice taking and clinical judgement: impact of advice distance and algorithm information. *Cognitive research: principles and implications*.

Rabin, M. (1994). A Model of Pre-game Communication. *Journal of Economic Theory*, 63(2):370–391.

Riboni, A. (2010). Committees as substitutes for commitment. *International Economic Review*, 51(1):213–236.

Salant, D. J. (1991). A repeated game with finitely lived overlapping generations of players. *Games and Economic Behavior*, 3(2):244–259.

Saragih, M. and Morrison, B. W. (2022). The effect of past algorithmic performance and decision significance on algorithmic advice acceptance. *International Journal of Human–Computer Interaction*, 38(13):1228–1237.

Schemmer, M., Hemmer, P., Kühl, N., Benz, C., and Satzger, G. (2022). Should i follow ai-based advice? measuring appropriate reliance in human-ai decision-making. *arXiv preprint arXiv:2204.06916*.

Schoonderwoerd, T. A., van Zoelen, E. M., van den Bosch, K., and Neerincx, M. A. (2022). Design patterns for human-ai co-learning: A wizard-of-oz evaluation in an urban-search-and-rescue task. *International Journal of Human-Computer Studies*, 164:102831.

Schultze, T., Rakotoarisoa, A.-F., and Schulz-Hardt, S. (2015). Effects of distance between initial estimates and advice on advice utilization. *Judgment & Decision Making*, 10(2).

Schulz, E. and Gershman, S. J. (2019). The algorithmic architecture of exploration in the human brain. *Current opinion in neurobiology*, 55:7–14.

Sekiguchi, T. (1997). Efficiency in repeated prisoner's dilemma with private monitoring. *Journal of Economic Theory*, 76:345–361.

Shani, G. and Gunawardana, A. (2011). Evaluating recommendation systems. In *Recommender systems handbook*, pages 257–297. Springer.

Sibert, A. (2003). Monetary policy committees: Individual and collective reputation. *Review of Economic Studies*, 70(3):649–665.

Smith, L. (1992). Folk theorems in overlapping generations games. *Games and Economic Behavior*, 4(3):426–449.

Soll, J. B., Palley, A. B., and Rader, C. A. (2022). The bad thing about good advice: Understanding when and how advice exacerbates overconfidence. *Management Science*, 68(4):2949–2969.

Speekenbrink, M. (2022). Chasing unknown bandits: Uncertainty guidance in learning and decision making. *Current Directions in Psychological Science*, 31(5):419–427.

Speekenbrink, M. and Konstantinidis, E. (2015). Uncertainty and exploration in a restless bandit problem. *Topics in cognitive science*, 7(2):351–367.

Stan Development Team (2022). RStan: the R interface to Stan. R package version 2.21.5.

Stigler, G. J. (1964). A Theory of Oligopoly. *Journal of Political Economy*, 72(1):44–61.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Tabellini, G. (2010). Culture and institutions: Economic development in the regions of Europe. *Journal of the European Economic Association*, 8(4):677–716.

Tauchert, C. and Mesbah, N. (2019). Following the robot? investigating users' utilization of advice from robo advisors.

Tauni, M. Z., Memon, Z. A., Fang, H.-X., Jebran, K., and Ahsan, T. (2019). Influence of investor and advisor big five personality congruence on futures trading behavior. *Emerging Markets Finance and Trade*, 55(15):3615–3630.

Thompson, W. R. (1933). Biometrika trust on the likelihood that one unknown probability exceeds another in view of the evidence of two samples.

Uotila, J., Maula, M., Keil, T., and Zahra, S. A. (2009). Exploration, exploitation, and financial performance: Analysis of s&p 500 corporations. *Strategic management journal*, 30(2):221–231.

Van Bruggen, G. H., Smidts, A., and Wierenga, B. (1998). Improving decision making by means of a marketing decision support system. *Management Science*, 44(5):645–658.

Van Swol, L. M., Paik, J. E., and Prahl, A. (2018). Advice recipients: The psychology of advice utilization.

Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and computing*, 27(5):1413–1432.

von Helversen, B., Mata, R., Samanez-Larkin, G. R., and Wilke, A. (2018). Foraging, exploration, or search? on the (lack of) convergent validity between three behavioral paradigms. *Evolutionary Behavioral Sciences*, 12(3):152.

Wang, Q., Huang, Y., Jasin, S., and Singh, P. V. (2022). Algorithmic transparency with strategic users. *Management Science*.

Wiehler, A., Chakroun, K., and Peters, J. (2021). Attenuated directed exploration during reinforcement learning in gambling disorder. *Journal of Neuroscience*, 41(11):2512–2522.

Wilson, R. C., Bonawitz, E., Costa, V. D., and Ebitz, R. B. (2021). Balancing exploration and exploitation with information and randomization. *Current opinion in behavioral sciences*, 38:49–56.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., and Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, 143(6):2074.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, Princeton, New Jersey.

Wu, C. M., Schulz, E., Pleskac, T. J., and Speekenbrink, M. (2022). Time pressure changes how people explore and respond to uncertainty. *Scientific reports*, 12(1):1–14.

Xu, X. and Potters, J. (2018). An experiment on cooperation in ongoing organizations. *Journal of Economic Behavior & Organization*, 147:28–40.

Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational behavior and human decision processes*, 93(1):1–13.

Yaniv, I. and Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational behavior and human decision processes*, 83(2):260–281.

Zajkowski, W. K., Kossut, M., and Wilson, R. C. (2017). A causal role for right frontopolar cortex in directed, but not random, exploration. *Elife*, 6:e27430.

Zak, P. J. and Knack, S. (2001). Trust and growth. *The economic journal*, 111(470):295–321.

Zettler, I., Thielmann, I., Hilbig, B. E., and Moshagen, M. (2020). The nomological net of the hexaco model of personality: A large-scale meta-analytic investigation. *Perspectives on Psychological Science*, 15(3):723–760.

Zhou, W., Lin, M., Xiao, M., and Fang, L. (2021). Exploitation and exploration: Improving search precision on e-commerce platforms. *Available at SSRN 3762144*.

Appendix A

A.1 Results From Post-Experiment Questions

In this section, we analyze the answers from the post-experiment questions and their relationship to the other behavior measurements. The summary of descriptive statistics is shown in Table A.1.1. We analyze how the reward earned from the participants and the number of switches between boxes correlates with the self-reported exploration and exploitation tendencies in Table A.1.2. Table A.1.3 and A.1.4 report the regressions of the explorative and the exploitative tendencies on the number of switches and the reward.

Table A.1.1: Summary of Post-Experiment Questions

treatment	explorative	exploitive	algorithm usefulness	algorithm relying	algorithm following	algorithm rating
control	.78(0.22)	.88(0.18)				
exploration	.81(0.16)	.90(0.11)	.50(0.31)	.45(0.31)	.54(0.29)	.45(0.29)
exploitation	.74(0.22)	.87(0.15)	.69(0.21)	.65(0.24)	.66(0.25)	.67(0.21)
balanced	.82(0.19)	.90(0.13)	.67(0.25)	.62(0.28)	.67(0.26)	.65(0.25)

Note: The descriptive statistics of the post-experiment questions for each treatment. The numbers outside brackets are the mean of the reported value, and the numbers within brackets are the standard deviation. The algorithm rating is calculated from the average of three algorithm ratings.

As the ratings on the algorithm were highly correlated, the three variables have been combined into a single variable, *algorithm rating*. The rating is significantly different between treatments ($F(2, 392) = 18.40$, $p < .001$, partial η^2=.086).

Table A.1.2: Correlations between Reward, Switch, and Explorative and Exploitative Tendency

	reward	switch	exploitative
switch	-.51(.26)		
explorative	.08(.01)	.07(.01)	
exploitative	.17(.03)	.11(.01)	.55(.30)

Note: The correlation between the explorative and exploitative tendency from the post-experiment questions and the number of switches and the amount of reward across the experiment. The numbers outside brackets are the correlation between variables, and the numbers within brackets are the R^2.

Table A.1.3: Regression Table of Post-Experiment Explorative and Exploitative Tendency on Reward

	Coefficient	SE	t-value	p-value
Intercept	53.08	2.16	24.48	.000
explorative	6.02	1.34	4.45	.000
exploitative	-2.73	0.97	-2.82	.005

Table A.1.4: Regression Table of Post-Experiment Explorative and Exploitative Tendency on Switch

	Coefficient	SE	t-value	p-value
Intercept	0.57	0.14	4.21	.000
explorative	-0.34	0.08	-3.99	.000
exploitative	0.22	0.06	3.68	.000

Table A.1.5: Correlations between Reported Algorithm Usefulness, Relying, Following, Combined Rating, and Estimated ρ

	algorithm usefulness	algorithm relying	algorithm following	algorithm rating
algorithm relying	.84(.71)			
algorithm following	.72(.52)	.82(.67)		
algorithm rating	.92(.85)	.96(.92)	.91(.82)	
ρ	.52(.27)	.60(.36)	.52(.27)	.59(.35)

Note: The correlation between the algorithm usefulness, relying, following, and rating from the post-experiment questions and the estimated ρ. The numbers outside brackets are the correlation between variables, and the numbers within brackets are the R^2.

Table A.1.6: Tukey HSD on Rating of Algorithm

treatments		mean difference	p-value	95% lower bound	95% upper bound
exploration	exploitation	-0.17	.000	-0.24	-0.10
	balanced	-0.15	.000	-0.22	-0.08
exploitation	balanced	0.02	.813	-0.05	0.09

A.2 Behavioral Results

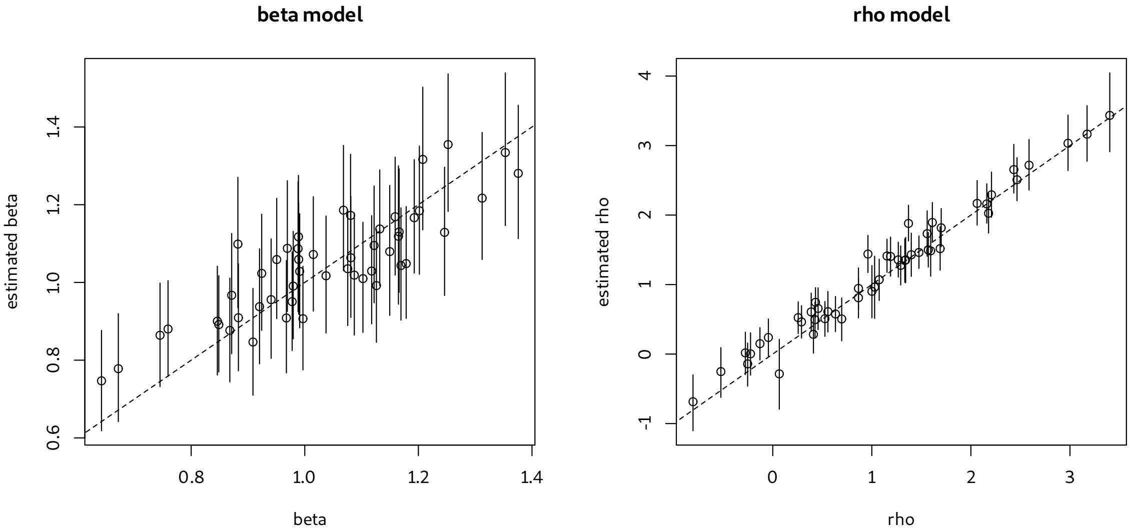
Table A.2.1: Regressions of Matching on the Difference in the Exploration Tendency between the Participants and the Algorithm

	all			exploitation			balanced			exploration		
	Est	SE	p-value	Est	SE	p-value	Est	SE	p-value	Est	SE	p-value
constant	131.70	4.93	$< .001$	173.20	9.81	$< .001$	155.90	4.53	$< .001$	98.14	13.77	$< .001$
difference	0.38	0.02	$< .001$	0.23	0.07	.002	0.41	0.04	$< .001$	0.43	0.07	$< .001$
difference^2	0.00	0.00	$< .001$	0.00	0.00	.097	0.00	0.00	$< .001$	0.00	0.00	.024
exploitation	25.10	5.17	$< .001$	-	-	-	-	-	-	-	-	-
balanced	22.77	4.91	$< .001$	-	-	-	-	-	-	-	-	-
N	442			146			148			148		
R^2	0.55			0.09			0.40			0.46		

Note: Linear regressions of number of matching in phase 2 on the difference in the exploration tendency between the participants and the algorithm. Difference in the exploration tendency is the number of switches by a participant in phase 1 minus the number of recommended switches in phase 2 from the algorithm. The variable difference^2 is squared difference. The "al" regression uses all data from the non-control treatments and the other three regressions use data from individual treatments. Treatment exploration is the base group for treatments exploration, exploitation and balanced. All models also include age and gender as socio-demographic control variables.

A.3 Parameter Recovery

Figure A.3.1: Parameter Recovery



Note: Results of a parameter recovery simulation for the two models. Scatter plots of true individual parameters (x-axis) versus estimated individual parameters (y-axis). Whiskers indicate 95% highest density credibility intervals for each parameter estimate.

A.4 Supplementary Results: Cognitive Modeling Analyses

Table A.4.1: Summary of Fitted Models

Model	Treatment	Tasks	Mean	Std	min(neff)	max(Rhat)	div	max(tree)	LOO
β	control	1-3	0.58	0.26	1218	1.0044	0	4	-46355
β	control	4-6	0.89	0.26	994	1.0012	0	4	-37882
β	exploitation	1-3	0.47	0.30	960	1.0039	0	4	-47161
β	exploitation	4-6	1.32	0.23	1901	1.0024	0	4	-24789
ρ	exploitation	4-6	3.55	1.38	480	1.0123	0	6	-23273
β	balanced	1-3	0.61	0.28	1165	1.0042	0	4	-45252
β	balanced	4-6	1.17	0.18	2665	1.0004	0	4	-28524
ρ	balanced	4-6	2.22	1.44	330	1.0107	0	5	-24136
β	exploration	1-3	0.42	0.28	997	1.0049	0	4	-50562
β	exploration	4-6	0.89	0.19	2343	1.0009	0	4	-38641
ρ	exploration	4-6	1.97	1.98	162	1.0494	0	6	-29862

Note: Population mean and standard deviation of the model parameter for each fitted model. The remaining columns report the minimum effective sample size, maximum Rhat, number of divergent transitions and the maximum tree-depth for each model. LOO is an estimate for the out-of-sample prediction accuracy of a Bayesian Model (Vehtari et al., 2017).

Table A.4.2: Regression Models for Inclination to Follow Recommendations (ρ)

	all			exploitation			balanced			exploration		
	Est	SE	p-value	Est	SE	p-value	Est	SE	p-value	Est	SE	p-value
constant	0.03	0.32	.931	0.86	0.67	.202	0.21	0.29	.476	-1.63	0.90	.071
difference	0.29	0.31	.347	0.43	0.49	.384	-0.09	0.35	.796	0.36	0.68	.595
difference^2	1.30	0.12	$< .001$	1.47	0.18	$< .001$	0.78	0.14	$< .001$	1.50	0.27	$< .001$
low explore	0.49	0.25	.052	-	-	-	-	-	-	-	-	-
optimal	-0.34	0.25	.183	-	-	-	-	-	-	-	-	-
N	442			146			148			148		
R^2	0.56			0.69			0.50			0.47		

Note: Linear regression of the latent tendency to follow algorithm recommendations on the difference and the squared difference between the participant's exploration tendency and the algorithm's true exploration tendency $log(\beta_1)$–$log(\beta_{exo})$. Models for the data of all treatments with age and gender as socio-demographic control variables.

A.5 Post-hoc Tests

This section summarizes post-hoc pairwise comparisons between the treatments using Tukey HSD corrections in phase 2.

Participants are more explorative in the exploration and control treatments than in the exploitation and balanced treatments. Between the control and the exploration treatment, and between the exploitation and the balanced treatment, no significant difference could be detected. Regarding rewards, participants in the exploitation treatment earn on average more than subjects in the control treatment. The number of choices that match with the algorithm's recommendation is higher in the balanced and exploitation treatments than in the exploration treatment.

Table A.5.1: Tukey HSD on Average Number of Switches

treatments		mean difference	p-value	95% lower bound	95% upper bound
exploitation	control	-7.72	$< .001$	-11.93	-3.51
	balanced	-2.12	0.559	-6.31	2.06
	exploration	-11.39	$< .001$	-15.58	-7.20
balanced	control	-5.59	0.004	-9.79	-1.40
	exploitation	2.21	0.559	-2.06	6.31
	exploration	-9.27	$< .001$	-13.44	-5.09
exploration	control	3.68	0.110	-0.52	7.87
	exploitation	11.39	$< .001$	3.51	11.93
	balanced	9.27	$< .001$	5.09	13.44

Table A.5.2: Tukey HSD on Average Rewards

treatments		mean difference	p-value	95% lower bound	95% upper bound
exploitation	control	1.58	$< .001$	0.38	2.77
	balanced	-0.41	0.814	-1.59	0.78
	exploration	0.46	0.749	-0.73	1.65
balanced	control	1.98	0.004	0.79	3.17
	exploitation	0.41	0.814	-0.78	1.59
	exploration	0.87	0.233	-0.31	2.05
exploration	control	1.12	0.074	-0.07	2.30
	exploitation	-0.46	0.749	-1.65	0.73
	balanced	-0.87	0.233	-2.05	0.31

Table A.5.3: Tukey HSD on Number of Matches

treatments		mean difference	p-value	95% lower bound	95% upper bound
exploitation	balanced	2.98	0.252	-1.44	7.40
	exploration	22.17	$< .001$	17.75	26.59
balanced	exploitation	-2.98	0.252	-7.40	1.44
	exploration	19.19	$< .001$	14.78	23.59

A.6 Changes Over Tasks

Figure A.6.1 shows changes over tasks. In general, average switches and entropy decrease over tasks and average rewards increase over tasks. The exploitation treatment and the balanced treatment have a higher decrease in switches in comparison to the other two treatments.

Figure A.6.1: Average Number of Switches, Average Rewards and Entropy over Tasks

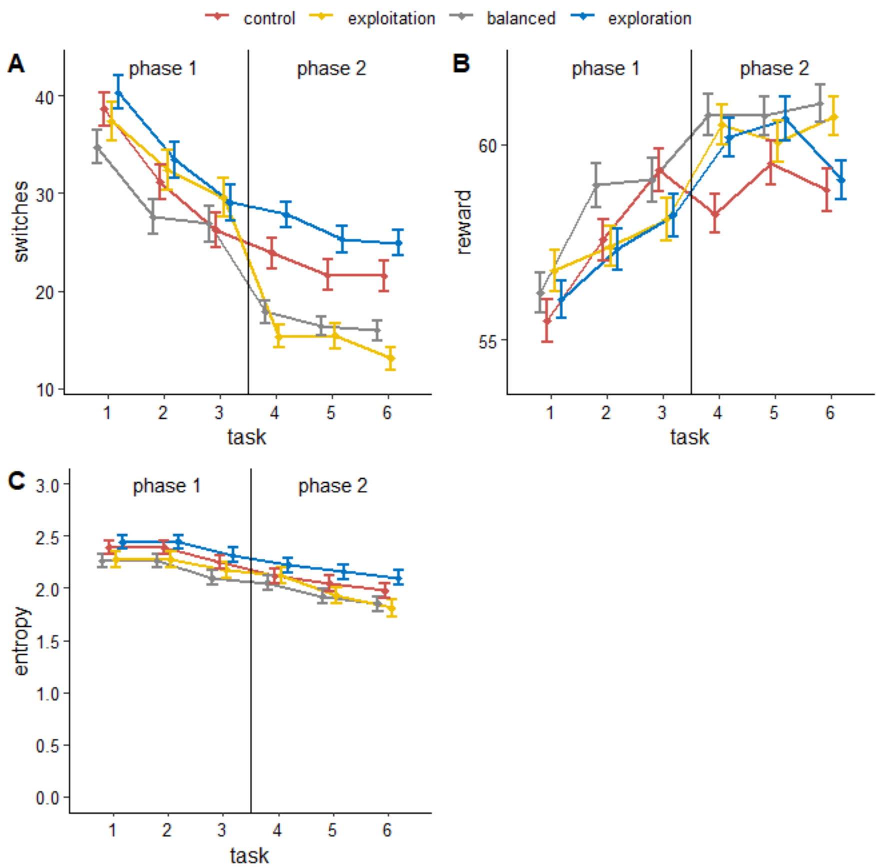
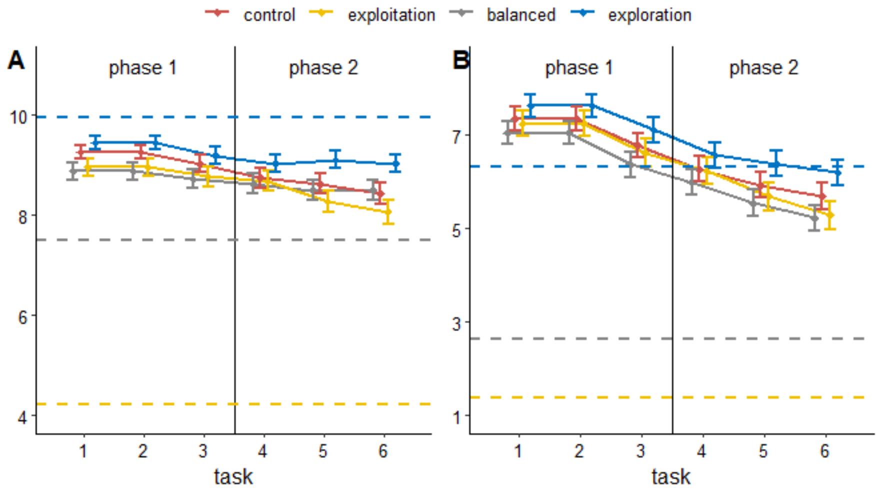


Figure A.6.2 shows how the average number of bandits explored changes over tasks. Panel A includes all trials, while Panel B includes the last 50 trials when participants' behavior is stabilized. We simulate with the three exploration parameters $\phi = 0.4$, $\phi = 1.4$ and $\phi = 3.3$ and calculate the average number of bandits that the algorithm would explore when it is exploitative, balanced, and explorative. The dotted lines in the figure are the simulation results. For all trials, participants explore on average more bandits than a balanced algorithm and an exploitative algorithm, but they explore less than an explorative algorithm. Looking at the last 50 trials, participants are more explorative than an explorative algorithm in phase 1. Participants in treatments other than exploration explore fewer bandits than an explorative algorithm in phase 2.

Figure A.6.2: Average Number of Bandits Explored.

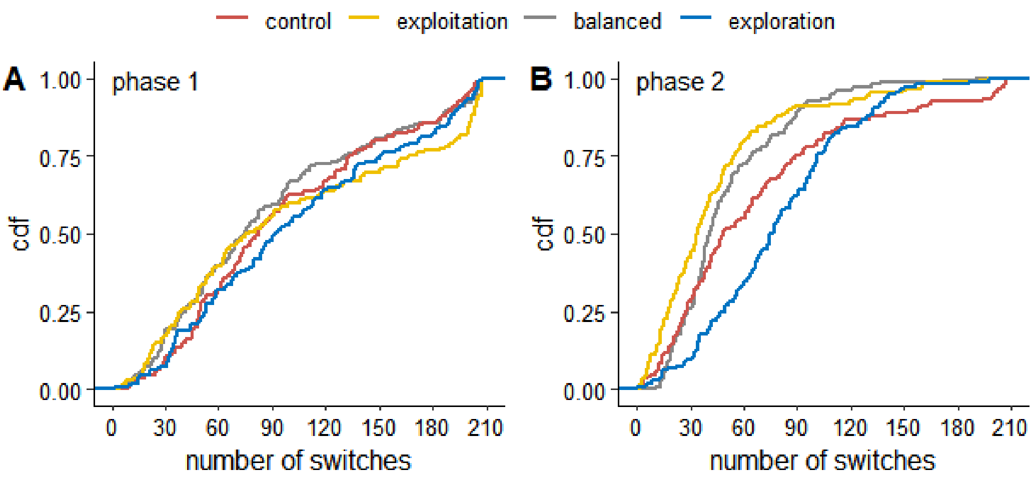


Note: Panel A includes all the trials. Panel B includes only the last 50 trials. The dotted lines are the average number of bandits explored by algorithm based on simulation.

A.7 Empirical CDF

The two panels in Figure A.7.1 are empirical CDF for all treatments in phase 1 and phase 2. There is not much difference in phase 1. In phase 2, participants in the exploration and control treatments are less explorative than those in the exploitation and balanced treatments. The control treatment first order statistically dominates the exploitation treatment.

Figure A.7.1: Empirical Cumulative Distribution Curve



A.8 Instructions for the Experiment

A.8.1 Consent Form

Dear Participant:

This is an academic study undertaken by the Faculty of Psychology at the University of Bremen, Germany. It seeks to advance knowledge about how people make decisions and how they learn to decide. Participation in this study is entirely voluntary, and you may cease participation at any time. If you decide to participate in this study, you must affirm that you understand the terms below and consent to participate. However, receiving compensation requires that you complete the experiment in its entirety.

The study takes about 25 minutes on average. You will see ten boxes on the screen and have to pick one of the boxes. Each box contains a different number of points. Your goal is to maximise the number of accumulated points. As compensation, you will receive a base payment of £2.5 plus an additional payment depending on how many points you accumulate through the choices you make. The amount of this additional payment can be between £0 and £11. However, receiving £11 is very unlikely. If you always choose the best or the second-best option, you will in most cases earn a bonus of about £3. At the beginning of this study, you will be asked to enter your Prolific ID, which is required to receive the payment. Except for your Prolific ID, no information will be tracked or stored that could be used to reveal your identity. The decisions you make and your answers during the study are pseudonymised. After you receive the payment, we will delete your Prolific ID from the data. From that point forward, we will no longer be able to identify your data or delete your data on request. We will use the collected data in the course of academic communication, which includes making it publicly available on the platform of the Open Science Framework (osf.io). There are no known physical, mental, social, or legal risks involved in the study. We kindly ask that you give the tasks your full attention and exit the full-screen mode only after completing the study. Please make sure that your participation is free from any distractions (e.g., mobile phone, television, music).

When you participate, you affirm that:

1. You will solve the tasks diligently and without the help of any tools.

2. You are over 18 years of age.

3. You understand that your participation is voluntary and that you may refuse to participate in this study or discontinue your participation at any point. Receiving the payment, however, is contingent on finishing the study.

4. You agree to the use of your data for the purpose of research and it being published anonymously for academic communication.

5. If you would like to receive a complete description and rationale for this study, or if you have questions regarding any concerns that arise from participating in this study, please direct your request to hslin@uni-bremen.de.

6. If you have any ethical concerns, please direct them to pbanik@uni-bremen.de.

If you agree with these terms, please select 'continue' to proceed.

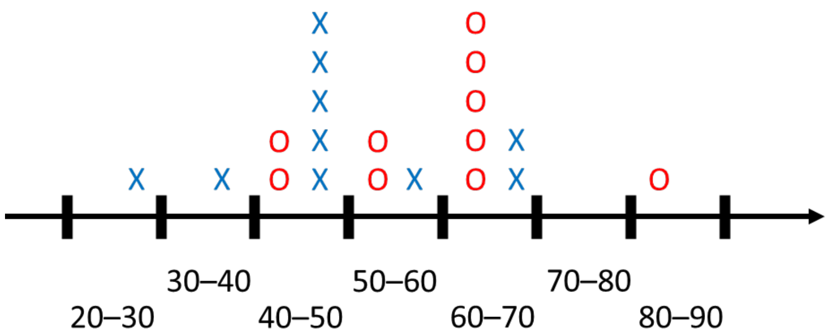
A.8.2 Instruction

To earn the highest possible amount during this study please carefully read this introductions page. Approximately half of your payment depends on your individual decisions you make later.

This experiment consists of two sessions, each consisting of a training block followed by three real decision-making blocks. In each block you will be asked to make repeated choices between ten boxes that are presented on the screen. In each instance you may choose one box to be opened by clicking on it. Once the box is chosen, a message below the box will show the number of points the box contained. Your goal is to maximise the number of points (i.e., the reward you receive from a box) over all decision-making blocks.

Each time you open a box, the reward that you receive may differ. Each box generates the reward from a fixed average in each block. This average reward is randomly selected from a normal distribution with a mean of 50 and standard deviation of 10. Each time a box is clicked, the reward is drawn from a normal distribution with the mean of the average reward of the clicked box and a standard deviation of 10. This means that, if you select a box with an average reward of 50 multiple times, the average of the generated reward will be close to 50. Half of the generated rewards will be between 43 and 57 points. For example, imagine you click ten times on a high-paying box (red circle, with an average reward of 60) and ten times on a low-paying box (blue cross, with an average reward of 45). The high-paying box may generate the rewards

60, 41, 81, 51, 69, 42, 54, 60, 60, 61, and the bad box may generate the rewards 48, 37, 49, 60, 29, 44, 65, 45, 48, 54. Thus, as the figure below illustrates, the low-paying box may sometimes result in rewards that are higher than the high-paying box. On average, however, you will earn a higher reward when clicking on the high-paying box.



During each block, the program will display the status of the current block on the top of the screen. It contains the number of choices you have made, the total amount of choices you can make in the current block, and the number of points you have earned in the current block.

Between each choice, you have unlimited time to decide which box you want to click on. Please carefully consider your choices.

For participating, you will receive a base payment of £2.5. In addition, you will receive a payment that depends on your choices during the experiment.

The additional payment is based on the average points in the boxes you opened across the six actual decision-making blocks. If you choose randomly and earn on average 45 points, you will receive no additional compensation. For every one point above 45, you will receive a bonus of 20 pence. For example, if you scored on average 55 points per trial, you will receive a bonus payment of £2. This bonus is capped at 100 points, i.e., you can receive a maximum of £11 in addition to the base payment of £2.5. However, it is unlikely that you reach a point average of 100 points. Typically, if you always choose the best or the second-best box available, you will likely reach a point average of about 60, which would result in a bonus payment of £3.

If you're ready, please select 'continue' to proceed.

A.8.3 Control Questions

To ensure you fully understand the instructions, please answer the following questions by selecting your answers with your mouse:

1. What's your goal in the task?

   a. Get as many points as possible.

   b. Explore as many boxes as possible.

   c. Choose a box as quickly as possible.

2. Which statement about the number of points you receive when choosing a box repeatedly is correct?

   a. The number of points in each box changes from trial to trial.

   b. Each box always gives the same number of points.

3. How should you achieve higher points in the experiment?

   a. By clicking the high-paying box.

   b. By clicking random boxes.

4. How is your payment determined?

   a. I receive only a base payment of £2.5.

   b. My payment depends on my choices. In addition to the base payment, I will receive a payment that depends on the average number of points I collect from the boxes.

Once you have finished answering the questions, please press 'continue' to confirm your answers.

A.8.4 Instruction Before First Training Block

The next block is a training block. Please familiarise yourself with the procedure. The points acquired during the training block do not count toward your final reward.

If you're ready for the training block, please press 'continue' to proceed.

A.8.5 Instruction After Training Block

The training block is over.

All the boxes will be reset to random states. You cannot carry over what you learned to the next phase.

If you're ready, please select 'continue' to proceed with the decision-making block.

A.8.6 Instruction After Each Testing Block

In this block, you acquired xxx points per trial!

All the boxes will be reset to random states. You cannot carry over what you learned to the next block.

If you're ready, please select 'continue' to proceed.

A.8.7 Instruction Before Second Training Block

If Participants Are In the Control Treatment.

The next session will be the same as the previous one. You will first see a training block and then three actual decision-making blocks. Please remember that the points acquired during the training block do not count toward your final reward.

If you're ready for the training block, please press 'continue' to proceed.

If Participants Are In the Algorithm Recommendation Treatments.

The next session follows the same general procedure as the previous session. However, in these blocks you will receive recommendations from a cutting-edge artificial intelligence (AI) algorithm, which will suggest which box to open next for each trial. The AI makes the recommendations based on the same information that you have. Thus, it knows about the general distribution of rewards in the boxes and it will learn to make better recommendations based on the information it receives by observing your choices within a block. The AI does not know the rewards from the unopened boxes.

The AI will suggest to you which box to open next by presenting a yellow square around the box. In each trial, you can choose whether you want to follow the suggestion from the AI or choose a box on your own. You will first be confronted with a training block and then three actual decision-making blocks. Please familiarise yourself with the procedure. The points acquired during the training block do not count toward your final reward.

If you understand the newly added AI and you're ready, please press 'continue' to proceed.

A.8.8 Instruction After the Second Phase

The testing is now over. Across the six decision-making blocks, you acquired xxx points per trial, which is converted to £yyy additional payment.

Before we finish the experiment, we would like to ask you a few questions.

If you're ready, please select 'continue' to proceed with the questions.

A.8.9 Questions

Effort Question.

Please give us your honest and subjective assessment of how much effort you put into maximising your profit. Your answer will NOT have any negative consequences. Please respond by clicking on the slider below.

Data Quality.

Is there any reason we should doubt your data quality (for example, because you were distracted or were not concentrating)? Please answer honestly; your answer will NOT affect the amount of reward you receive or have any negative consequence.

My data may be flawed.

My data is fine.

Exploitation Tendency.

How much did you try to keep updated about how good all the boxes are? Please respond by clicking on the slider below.

Exploration Tendency.

How much did you try to choose the box in each trial with the highest average profit? Please respond by clicking on the slider below.

Algorithm Usefulness.

How useful were the suggestions from the AI? Please respond by clicking on the slider below.

Relying On the Algorithm.

How much did you rely on the suggestion from the AI in general? Please respond by clicking on the slider below.

Following the Algorithm.

How much did you follow the suggestions from the AI when you were unsure about which box to choose next? Please respond by clicking on the slider below.

Further Comments.

If you have any comments or remarks about the study, you can leave them anonymously in the box. If not, you may leave the textbox empty. Once you have finished, please press 'continue' to proceed.

Appendix B

B.1 Proof of Theoretical Prediction 3

The first part of the prediction is easy to show. Consider the following strategy profile: The public never transfers resources to the organization. Irrespective of the history, members of the organization always vote in favor of not sending transfers to the public. It is obvious that such a strategy profile constitutes a subgame-perfect Nash equilibrium and that it satisfies our sincere-voting requirement.

To show that equilibria with full cooperation exist, consider the following strategies. Decision-makers in their final terms always vote against sending transfers to the public, conditional on the organization having received transfers from the public. The other decision-makers vote in favor of sending transfers to the public unless the public did not submit transfers in the previous round or, in the previous round, the organization did not submit transfers to the public. In these cases, all decision-makers vote against sending transfers. The public always sends transfers to the organization unless it itself or the organization did not send transfers in the previous round.

The behavior of the public is optimal, as it is payoff-maximizing to send transfers to the organization exactly in those periods where the organization is expected to reciprocate this behavior. The behavior of decision-makers in their last terms is optimal as well, as not sending transfers to the public maximizes the payoffs in the last period. The other decision-makers have to weigh the current gains from defecting, which are $E + R_1 - (E + R_1 - T_2)$, against the discounted forgone future gains from cooperating, i.e. $\delta\,(E + R_1 - T_2 - E)$. The assumed behavior is optimal if $E + R_1 - (E + R_1 - T_2) \leq \delta (E + R_1 - T_2 - E)$, which is equivalent to (2.1).

It may be interesting to discuss strategies where defections lead to more extended periods of punishments. Suppose, for example, cooperation would break down for two periods following a defection, i.e. a period where the public or the organization do not send transfers. Could cooperation be sustained for larger sets of parameters than the one characterized by (2.1)? Actually, this is not the case because the middle-aged decision-maker faces only one additional period. As a consequence, stricter punishments like a two-period break down of cooperation after defection or even grim-trigger strategies do not increase the scope for cooperation. This implies that cooperation cannot occur if (2.1) is violated. □

B.2 Simulations for Assessing the Statistical Power

In the simulations, we iterated the following process 20,000 times for various effect sizes:

1. Create a data set of 5 (supergames) * 3 (number of matching groups per treatment) * 3 (number of groups per matching groups) * 2 (treatments) observations and an indicator variable for treatment 2.

2. Draw random numbers from the Bernoulli distribution with success probability 0.4 for treatment 1.

3. Draw random numbers from the Bernoulli distribution with success probability 0.4 + (effect size) for treatment 2.

4. Average the random draws within each group. These are the simulated average cooperation rates.

5. Regress the average cooperation rates on the indicator variable to get the difference and the cluster-adjusted standard error (clustering on the matching group). Use these for a one-sided $t-$test. Return the $p-$value.

Finally, we computed the share of the $p-$values smaller than 0.05, which gave us the statistical power.

Our simulations suggest that we will have enough power ($> 87\%$) to detect effect sizes of 15 percentage points at the 5% level with a one-sided $t-$test and 6 matching groups per treatment. We checked for the accuracy of the method by running a simulation with an effect size of 0 (again with 20,000 iterations of the process described above). The relative frequency of $p-$values smaller 0.05 was 0.051, which is very close to the 0.05 that we would expect in this case.

B.3 Probability of Playing "Send" After Memory-one Histories

For the decision-makers, we assume that they condition their choices on the public's choice, the committee's choice, and their own choice in the previous round. The possible memory-one histories for M1 differ from those for M2 and M3 because M1 has just entered the committee and thus does not have a history of his/her own choice in the previous round. For M1 of all treatments, the memory-one histories are ($\emptyset$, ss, sn, ns, nn), where the first letter denotes the choice of the public in the previous round and the second letter denotes the choice of the committee in the previous round. These histories further imply that the choices made by the committee when M1 is not in the committee affect the choices of M1. For M2 and M3, the memory-one histories are ($\emptyset$, sss, ssn, sns. snn, nss, nsn, nns, nnn). The first letter denotes the choice of the public, the second letter is the choice of the committee, and the third letter is one's own choice in the previous round. Notice that for treatment I, the committee's choice is the decision-maker's choice; the only possible histories are (sss, snn, nss, nnn). Notice that for treatment I and CST, M2 and M3 do not have empty history because they made decisions as M1 and M2 in the previous round. Table B.3.1 shows the probability of playing send for M1, M2 and M3 for each treatment.

The probability for M1 to play "Send" is high irrespective of the choices made either by the public or buy the committee in the last round. Overall, M3 sends less often than M2 and M1.

Table B.3.1: Cooperation rates after memory-one histories for M1, M2 and M3

	$\sigma_\emptyset$	σ_{ss}	σ_{sn}	σ_{ns}	σ_{nn}	ln$L$
I	0.84	0.80	1.00	0.82	0.77	-44.61
	(0.08)	(0.10)	-	(0.15)	(0.16)	
C	0.72	0.77	0.68	0.66	0.72	-285.39
	(0.05)	(0.04)	(0.07)	(0.06)	(0.07)	
CST	0.68	0.71	0.64	0.75	0.78	-373.00
	(0.04)	(0.05)	(0.07)	(0.07)	(0.07)	
CA	0.75	0.77	0.64	0.64	0.38	-281.83
	(0.05)	(0.04)	(0.07)	(0.08)	(0.08)	

(1) M1

	$\sigma_\emptyset$	σ_{sss}	σ_{ssn}	σ_{sns}	σ_{snn}	σ_{nss}	σ_{nsn}	σ_{nns}	σ_{nnn}	ln$L$
I	-	0.91	-	-	0.44	0.95	-	-	0.50	-26.61
		(0.08)			(0.32)	(0.05)			(0.45)	
C	0.65	0.92	0.20	0.86	0.14	0.80	0.27	0.90	0.26	-217.42
	(0.06)	(0.03)	(0.08)	(0.10)	(0.06)	(0.05)	(0.15)	(0.05)	(0.08)	
CST	-	0.88	0.28	0.77	0.28	0.82	0.25	0.71	0.18	-224.74
		(0.03)	(0.07)	(0.11)	(0.06)	(0.05)	(0.10)	(0.19)	(0.09)	
CA	0.75	0.88	0.26	0.88	0.17	0.82	0.44	0.85	0.31	-226.08
	(0.05)	(0.03)	(0.09)	(0.08)	(0.06)	(0.06)	(0.18)	(0.11)	(0.09)	

(2) M2

	$\sigma_\emptyset$	σ_{sss}	σ_{ssn}	σ_{sns}	σ_{snn}	σ_{nss}	σ_{nsn}	σ_{nns}	σ_{nnn}	ln$L$
I	-	0.76	-	-	0.50	0.54	-	-	0.50	-36.43
		(0.10)			(0.90)	(0.16)			(0.36)	
C	0.51	0.76	0.17	0.69	0.20	0.72	0.19	0.69	0.19	-277.95
	(0.06)	(0.05)	(0.08)	(0.14)	(0.06)	(0.07)	(0.10)	(0.14)	(0.07)	
CST	-	0.77	0.16	0.75	0.23	0.81	0.12	0.67	0.23	-189.31
		(0.04)	(0.06)	(0.12)	(0.08)	(0.08)	(0.13)	(0.12)	(0.07)	
CA	0.64	0.77	0.12	0.47	0.22	0.61	0.50	1.00	0.44	-275.81
	(0.05)	(0.04)	(0.06)	(0.15)	(0.06)	(0.09)	(0.24)	-	(0.09)	

(3) M3

Note: Estimated probabilities of playing "Send" after memory-one histories. The three sub-tables from top to bottom are estimated results for M1, M2 and M3 in all treatments. Numbers in parentheses are the standard errors. ln$L$ is the log-likelihood of the model.

B.4 Candidate Strategies

Table B.4.1: Strategies 1-9 for the public

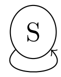
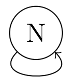
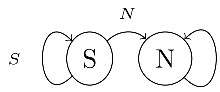
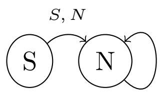
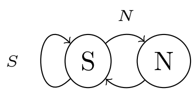
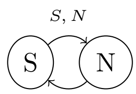
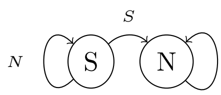
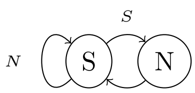
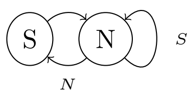
Acronym	Description	Automaton
ALLS	Always play “Send”.	
ALLN	Always play “Not Send”.	
GRIM	Play “Send” as long as the committee returns, otherwise play “Not Send” forever.	
FS	Play “Send” for one round, and irrespective of what the committee plays, play “Not Send” forever.	
TFT	Play “Send” as long as the committee returns, otherwise play “Not Send” for one round and return to “Send”.	
SN	Start with “Send”, then alternate between “Not Send” and “Send”.	
GRIMS	Play “Send” as long as the committee keeps, otherwise play “Not Send” forever.	
TFTS	Play “Send” as long as the committee keeps, otherwise play “Not Send” for one round and return to “Send”.	
NTFT	Start with “Not Send”, and play “Send” in the next round. Keep playing “Send” as long as the committee returns, otherwise go back to “Not Send”.	

Table B.4.2: Strategies 10-18 for the public

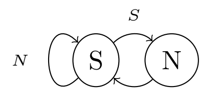
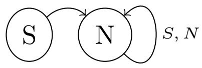
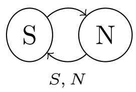
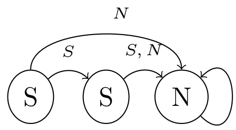
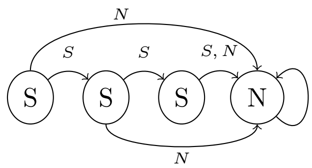
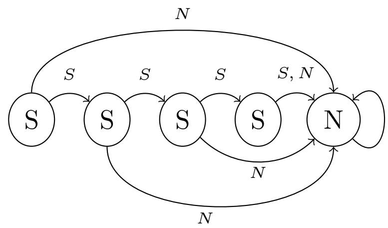
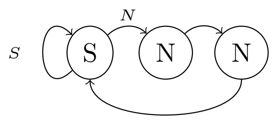
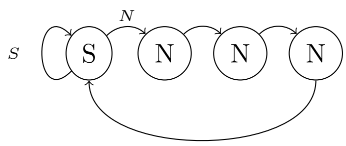
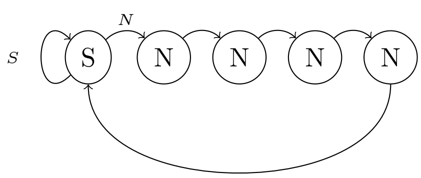
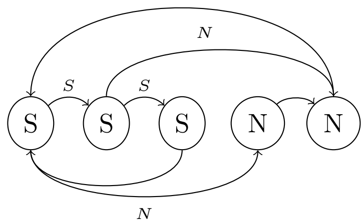
Acronym	Description	Automaton
NTFTS	Start with "Not Send", and play "Send" in the next round. Keep playing "Send" as long as the committee keeps, otherwise go back to "Not Send".	
FN	Start with "Not Send", then play "Send" forever.	
NS	Start with "Not Send", then alternate between "Send" and "Not Send".	
FS2	Play "Send" twice as long as the committee sends, then play "Not Send" forever.	
FS3	Play "Send" three times as long as the committee sends, then play "Not Send" forever.	
FS4	Play "Send" four times as long as the committee sends, then play "Not Send" forever.	
T2	Play "Send" as long as the committee returns, otherwise play "Not Send" for two rounds and return to "Send".	
T3	Play "Send" as long as the committee returns, otherwise play "Not Send" for three rounds and return to "Send".	
T4	Play "Send" as long as the committee returns, otherwise play "Not Send" for four rounds and return to "Send".	

Table B.4.3: Strategy 19 for the public

Acronym	Description	Automaton
T-MULT3	Play "Send" as long as the committee returns, otherwise play "Not Send" until the end of round 3 and round 6, then return to "Send".	

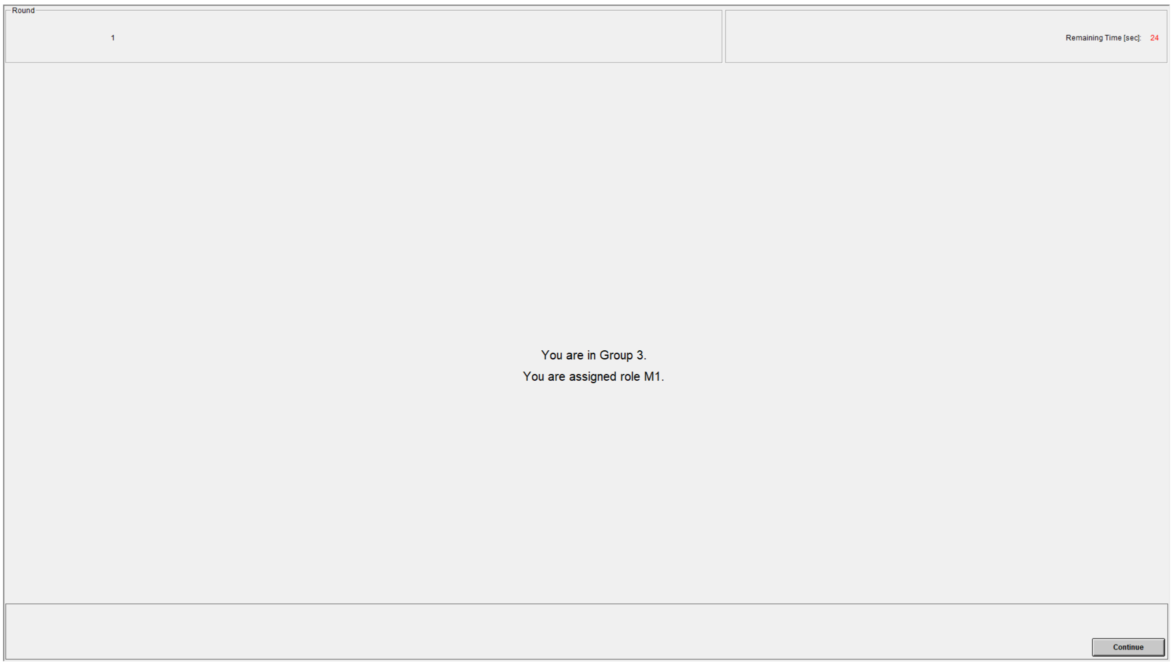
Notes: circles represent the states of an automaton. The first state from the left is the starting state. The labels S and N inside the nodes represent the choice of the public. Arrows represent deterministic state transitions. The labels on the arrow indicate the choices of the committee that trigger this transition. An unlabeled arrows indicates an unconditional transition that occurs independent of the observed profile.

B.5 Experimental Procedure

In this Section, we provide further information on the experimental procedure, including screen shots and the instructions for the C and the I treatment. The instructions for the other treatments are very similar and, therefore, we omit them here.

We ask subjects to read through instructions, and to answer quiz questions to make sure they understand before experiment starts. Once the experiment starts, subjects are informed about their role and group assignment in the current interaction on the first screen. Figure 3 is an example of what a subject would see that was assigned into Group 2 as a middle-aged decision-maker in treatment C.

Figure B.5.1: Assignment Stage



In the treatments without mission statements, the public and decision-makers decide between 'Send' and 'Not Send" which are represented by options A and B for the public, and options X and Y for the decision-makers. Figure 4 and 5 show the decision screens for the public and decision-makers, respectively in treatments C and CST.

Figure B.5.2: Choice Stage for Public

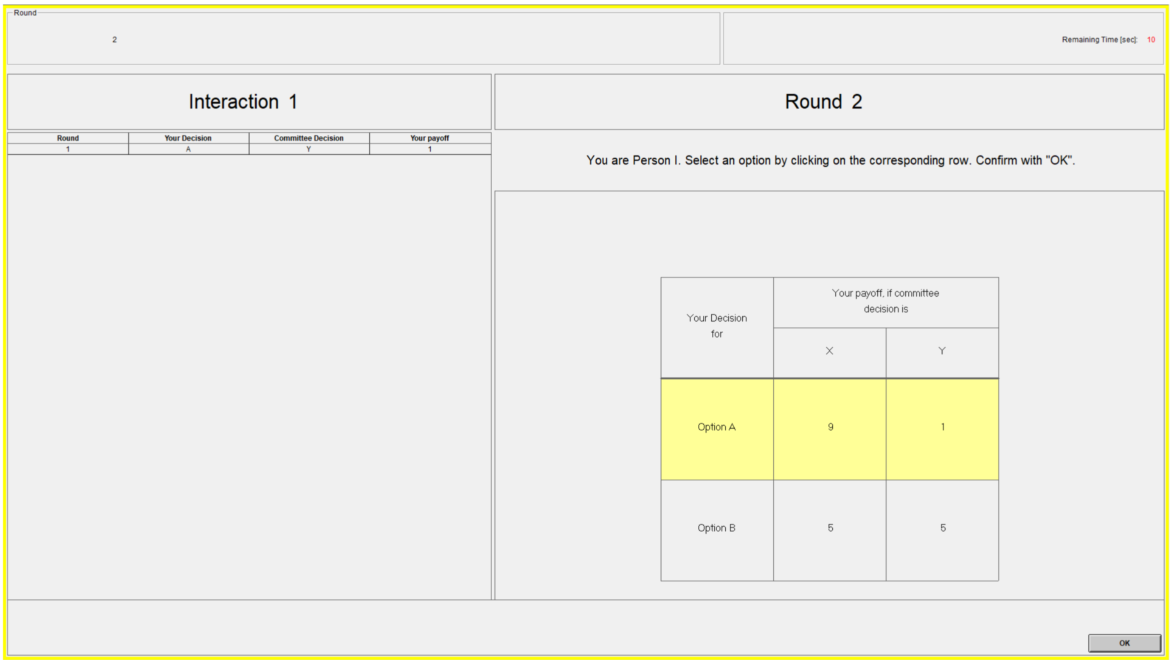
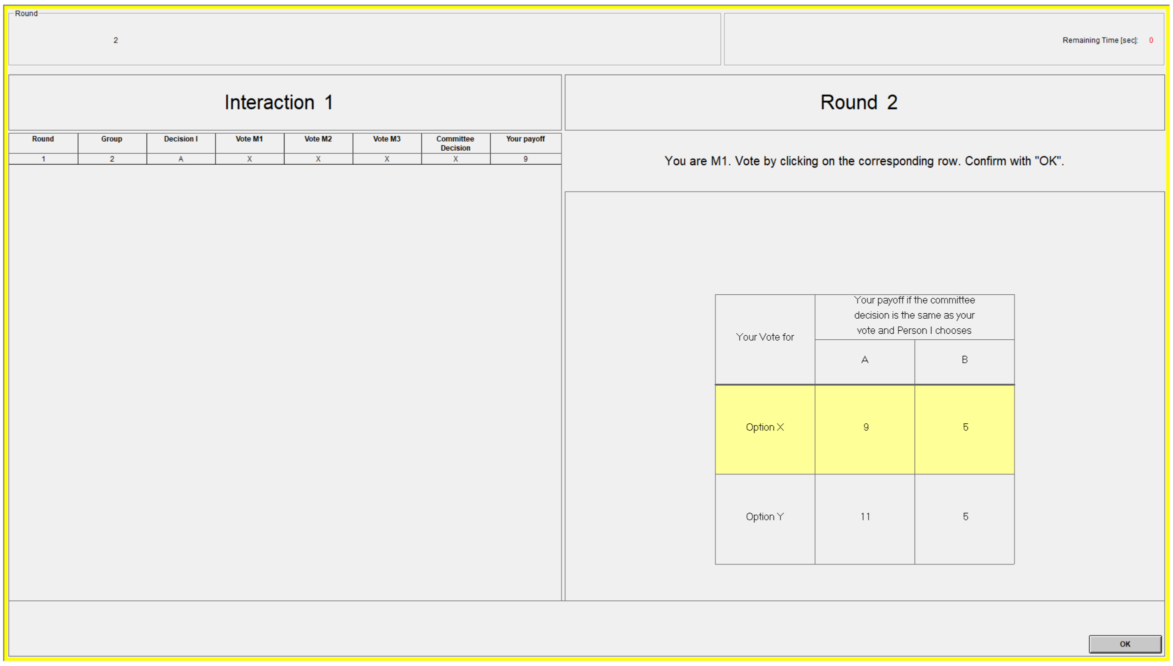
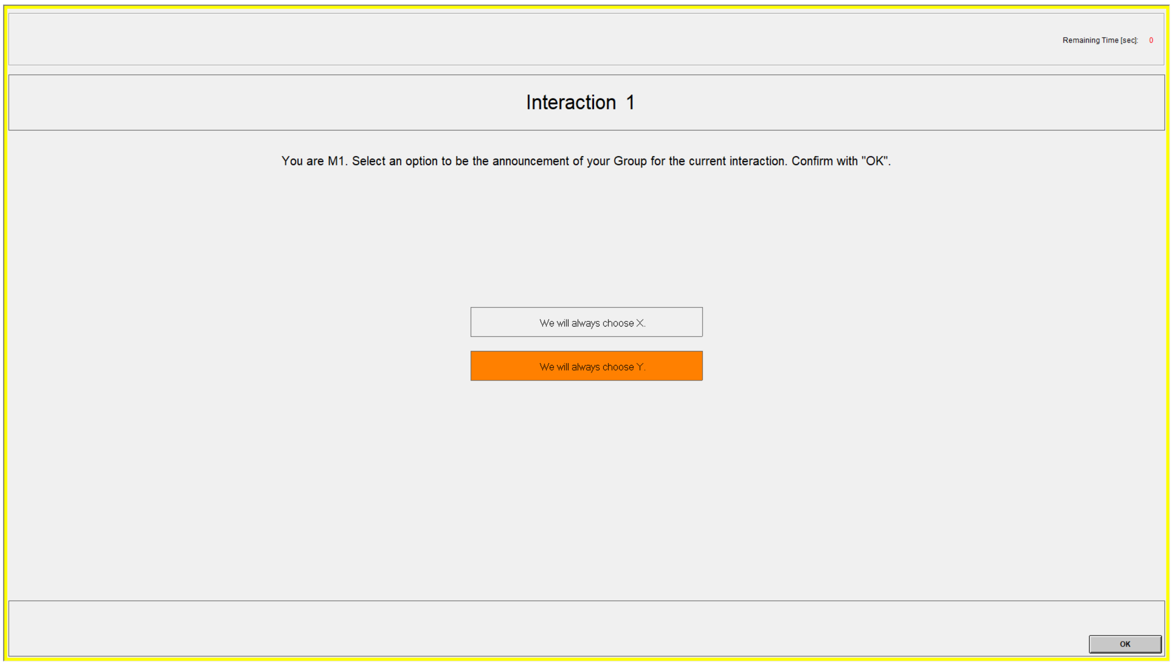


Figure B.5.3: Choice Stage for Decision-makers



In treatment CA, decision-makers engage in an additional voting stage on the announcement (group mission) before making their choices in the stage game. They vote between the two options "We will always choose X" and "We will always choose Y", as depicted in Figure 6.

Figure B.5.4: Announcement Stage for Decision-makers



The committee decision regarding the announcement will appear as "Annoucement" on screens of all subjects of that group in choice and feedback stages. Figure 7 gives an example of a choice screen with announcement. Figure 8 is a feedback stage that immediately follows the choice stage.

Between rounds, the continuation decision of the current supergame is represented by the rolling of a dice. As Figure 9 shows, the current interaction ends if the roll of the die is "1".

At the end of the session, we present subjects a payoff table with payment from all periods. Afterwards, subjects answer a standard socio-economic survey to elicit basic demographic information of the participants of our experiment.

Figure B.5.5: Choice Stage with Announcement

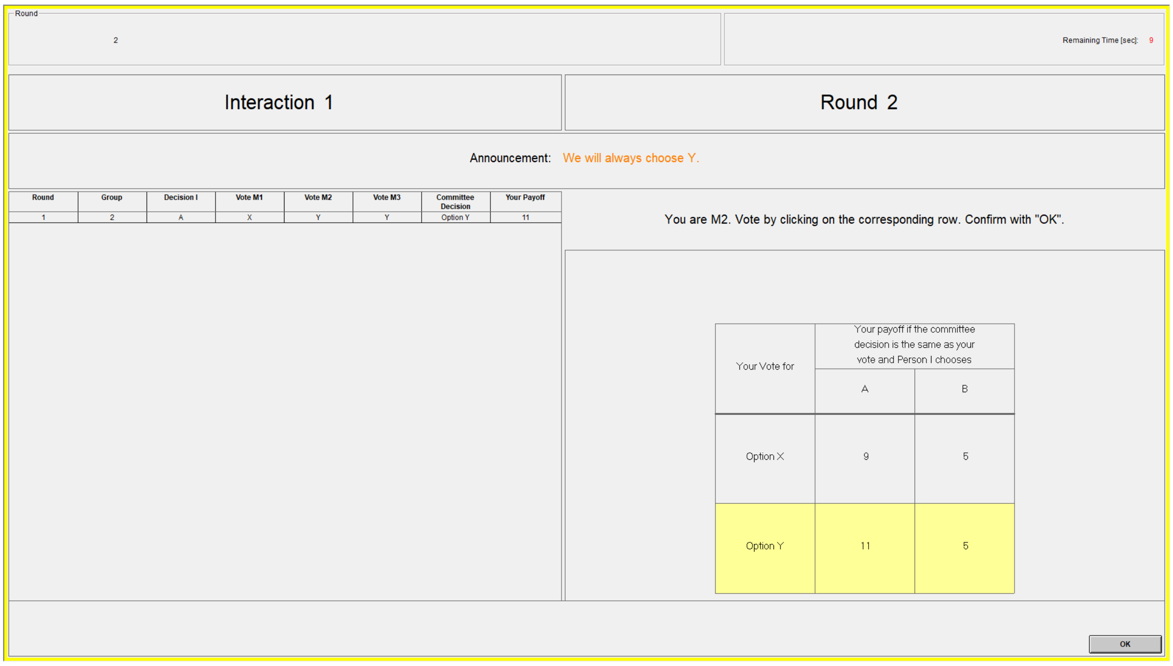


Figure B.5.6: Feedback Stage with Announcement

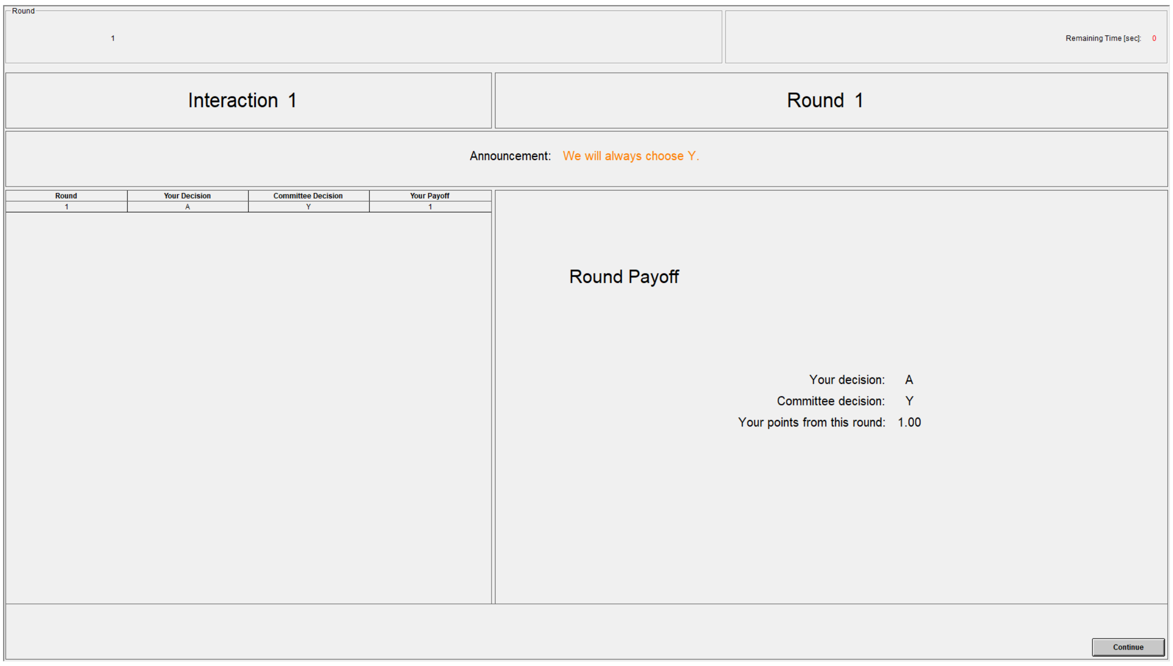
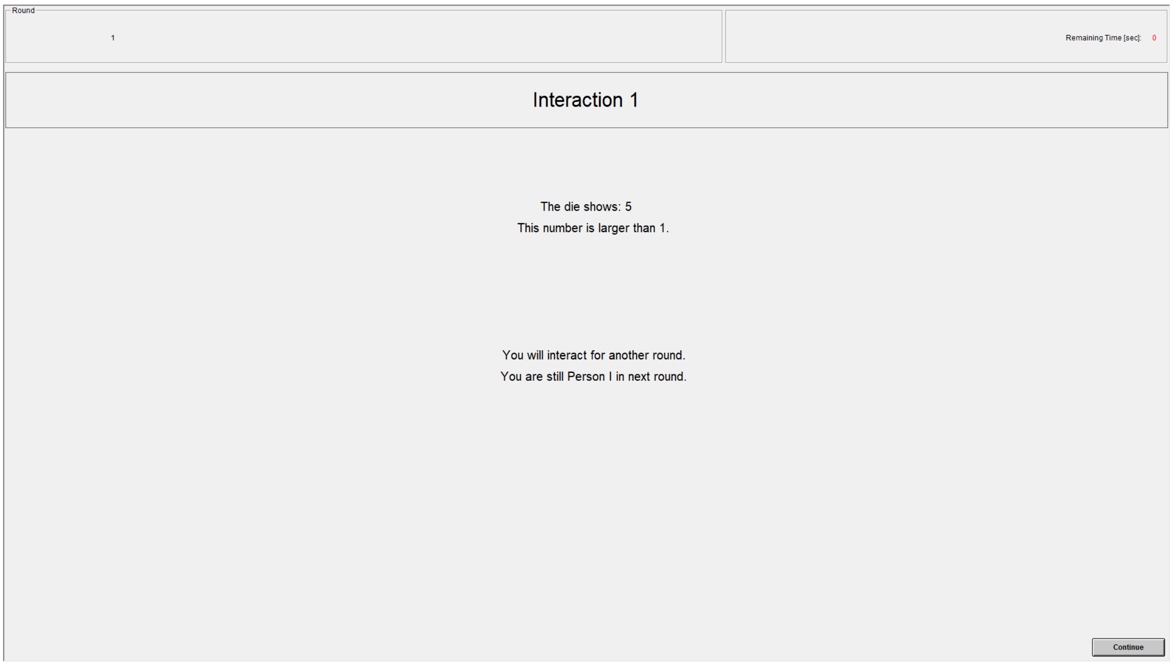


Figure B.5.7: Die Rolling for Continuation



[Instructions for Treatment C]

Overview

Welcome to this experiment. We please you not to talk with other participants during this experiment and to switch off your mobile devices.

You will be paid in cash for today's participation at the end of the experiment. The amount of money you receive depends on your own decisions, the other participants' decisions, and pure chance. It is important that you understand the instructions before the experiment starts.

In this experiment, every interaction between the participants runs through the computers you are sitting in front of. They will interact with each other anonymously. Neither your name nor the names of other participants will be announced. Also, for the evaluations only the anonymized data are used.

Today's session consists of several interactions, which typically consist of several rounds. Your payoff amount is the sum of all points earned, converted into euros, plus a **show-up fee of 5 EUR**. The points shall be converted into euro as follows. Every point is worth **5 cents**, so that:

20 Points = 1,00 EUR.

All participants are paid privately, so other participants cannot see how much you have earned.

Experiment

Interactions and role assignment

This experiment consists of several interactions, which are identical in their sequence. Each interaction consists of one or more rounds of committee meetings. The number of rounds of an interaction is random.

At any given time, there are three committees, with each committee consisting of three current members M1, M2 and M3. Each committee is assigned three potential waiting members W1, W2 and W3 and a person I who is not a committee member (I stands for "Individual").

At the beginning of every interaction all participants are randomly reassigned to the three committee groups (with three M, three W and one person I each). These committee groups exist for the duration of the interaction. In addition, the roles are randomly reassigned at the beginning of each interaction. The only exception is Role I. Whoever is assigned this role in the first interaction of the experiment will keep it throughout the entire experiment. Persons who are assigned a role other than person I at the beginning of the first interaction cannot be assigned role I later.

Both the membership in the committee and the possible waiting period is a maximum of three rounds. The number after the letter M indicates the term of office. After each round the term of office is increased by 1, i.e. M1 becomes M2 and M2 becomes M3. Person M3 leaves the committee at the end of the round and is assigned to another committee with the role W1. In the same way, the role of the persons in the waiting state is changed. W1 becomes W2 in the next round and W2 becomes W3. A person with W3 has finished the waiting time after the corresponding round and starts as a new member with the role M1. At any given time, there is exactly one role for each of the three committees.

Length of an interaction

The length of an interaction is random. After each round there is a 1/6 chance that the interaction ends, and all committees are dissolved. In this case, a new interaction starts with a new random allocation of roles (whereby - as already explained - participants in the role of person I always keep this role).

To determine the length of the interactions, a random number generator was used before the experiment to generate for each interaction a series of equally distributed random numbers from the set 1, 2, 3, 4, 5 and 6, as with the throws of an ordinary six-sided dice. It was determined after how many throws the number 1 appears for the first time. The number of these throws gives the length of the interaction.

Example: The random numbers 4, 6, 3, 4, 1 were generated. So, the number 1 appeared for the first time on the fifth throw. Thus, the interaction has the length 5.

If an interaction is very long, it is possible that a member who retires from a committee will become a member of the same committee again at a later date. The probability that the member M3 of a committee, who would thus retire in the next round, would ever re-enter that committee is about 5%. For the other members this probability is even lower.

Interaction und round schedule

In each round, person I decides between action A and action B by clicking on the corresponding row in the table on the screen (see Fig. 1, left panel). At the same time the committee members M1, M2 and M3 vote simultaneously for action X or Y by clicking on the corresponding row in the table on the screen (see Fig. 1, right panel). Abstentions are not possible. The option that receives more votes, is the decision of the committee. For example, if M1 votes for X, M2 for Y, and M3 for Y, the committee decision is Y.

Fig. 1: Decision masks of a person I (left) and a person M1, M2 or M3 (right)

Your Decision for	Your payoff, if committee decision is	
	X	Y
Option A	9	1
Option B	5	5

Your Vote for	Your payoff if the committee decision is the same as your vote and Person I chooses	
	A	B
Option X	9	5
Option Y	11	5

In case person I chooses option B, her payoff is 5 points and the payoff of every committee member is also 5 points, irrespective of the committee's decision. In case person I chooses option A, her payoff and the payoff of all committee members depends on the committee's decision. In case the committee decides for option X, person I receives 9 points and each committee member also receives 9 points. In case the committee decides for option Y, person I receives 1 point and each committee member receives 11 points.

The persons in the waiting state W1, W2 and W3 do not make a decision and receive a fixed payment of 5 points.

At the end of the round, all players assigned to a committee (M1, M2 and M3 and the waiting persons W1, W2, W3) receive information about the committee's collective decision and the individual votes of all committee members as well as person I's choice and the resulting payoff. Person I receives information about her payoff and in case she chose action A, she also learns the committee's decision (X or Y) but not the individual votes of the committee members. In case person I chose B, she does not receive information about the committee's decision

All rounds are identical in terms of the procedure. The progress of the current interaction is displayed in tabular form in each round for the committee to which you are currently assigned.

End and final payoff

As soon as chance ends the last interaction, the experiment is over.

At the end of the experiment all interactions are paid off. The total amount of points from all rounds will be converted into Euros and paid out privately.

On the last screen of the last round of the last interaction, you can see how much you have earned in Euros.

Questions?

If you have any questions, please contact us. An experimenter will then come to your place.

If you think you have understood everything well, you may start the quiz on the screen. This quiz is only to make sure that everyone has understood the instructions well. The answers will not affect your payoff.

[Instructions for Treatment I]

Overview

Welcome to this experiment. We please you not to talk with other participants during this experiment and to switch off your mobile devices.

You will be paid in cash for today's participation at the end of the experiment. The amount of money you receive depends on your own decisions, the other participants' decisions, and pure chance. It is important that you understand the instructions before the experiment starts.

In this experiment, every interaction between the participants runs through the computers you are sitting in front of. They will interact with each other anonymously. Neither your name nor the names of other participants will be announced. Also, for the evaluations only the anonymized data are used.

Today's session consists of several interactions, which typically consist of several rounds. Your payoff amount is the sum of all points earned, converted into euros, plus a **show-up fee of 5 EUR**. The points shall be converted into euro as follows. Every point is worth **5 cents**, so that:

20 Points = 1,00 EUR.

All participants are paid privately, so other participants cannot see how much you have earned.

Experiment

Interactions and role assignment

This experiment consists of several interactions that are identical in terms of their sequence. Each interaction consists of one or more rounds. The number of rounds of an interaction is random.

At the very beginning, before the first interaction, you and other participants are randomly assigned to a matching group of 9 people in total, which will remain for the entire experiment. You will only interact with the members of this matching group during the whole experiment.

At any given time, your matching group consists of three subgroups, each subgroup consisting of three people. Before each interaction, the members of your matching group are randomly distributed among the three subgroups. Of the three subgroup members, one person is the current decision maker (role E1, E2 or E3) and one person is waiting (roles W1, W2 or W3). The waiting person may replace the decision-maker in the future. The third person has the role person I. The letter I stands for "Individual". In addition, there are two passive assistants of the decision maker P1 and P2, who are randomly chosen from a different matching group, never make any choices in your matching group but receive payoffs at the very end of the experiment which depend on the choices made by E and I. You might also be selected to be such a passive assistant for a group in another matching group in addition to the role that you have in your own matching group but you will only learn about this at the very end of the experiment. In other words, there is no way in which you can influence decisions in other matching groups and your choices cannot be influenced by people outside your matching group either.

At the beginning of the first interaction you will be randomly assigned one of the roles E1, W1 or A. If you are assigned the role of person I, you will keep it throughout the experiment. If you are assigned the role E1 or W1, you will be randomly reassigned one of these roles at the beginning of each new interaction.

The maximum term of office for decision-makers is three periods. The number after the letter E indicates the term of office for the decision-maker. After each round, the term of office is increased by 1, that is, E1 becomes E2 and E2 becomes E3. A person with role E3 leaves the subgroup at the end of the round and is assigned to another subgroup with role W1. In the same way, the role of persons in the waiting state is changed. W1 becomes W2 and W2 becomes W3 in the next round. A person with W3 has finished the waiting time after the end of the corresponding round and starts as a new decision maker of his or her subgroup with the role E1 (unless the interaction ends after this round).

Length of an interaction

The length of an interaction is random. After each round there is a 1/6 probability that the interaction ends, and all subgroups of the matching groups are dissolved. In this case, a new interaction starts with a new random allocation of roles (whereby - as already explained - participants with the role of person I always keep this role and can only meet the other 8 people in your matching group).

To determine the length of the interactions, a random number generator was used before the experiment to generate for each interaction a series of equally distributed random numbers from the set 1, 2, 3, 4, 5 and 6, as with the throws of an ordinary six-sided dice. It was determined after how many throws the number 1 appears for the first time. The number of these throws gives the length of the interaction.

Example: The random numbers 4, 6, 3, 4, 1 were generated. So, the number 1 appeared for the first time on the fifth throw. Thus, the interaction has the length 5.

If an interaction is very long, it is possible that a decision-maker who leaves a subgroup will later become a decision-maker again in the same subgroup with the same person I. The probability that a decision maker E3, who would thus leave in the next round, would meet the same person again during this interaction is about 5%. This probability is even lower for decision-makers with a lower term of office, i.e. for E1 and E2.

Interaction and round schedule

In each round, person I decides between action A and action B by clicking on the corresponding row in the table on the screen (see Fig. 1, left panel). At the same time the decision-maker E (E1, E2 or E3) chooses action X or Y by clicking on the corresponding row in the table on the screen (see Fig. 1, right panel).

In case person I chooses option B, her payoff is 5 points and the payoff of E and her passive assistants P1 and P2 is also 5 points each, irrespective of E's decision. In case person I chooses option A, her payoff and the payoff of E, P1 and P2 depend on the E's decision. In case E decides for option X, person I receives 9 points and E, P1 and P2 also receive 9 points each. In case E decides for option Y, person I receives 1 point and E, P1 and P2 receive 11 points.

The person in the waiting state (W1, W2 or W3) does not make a decision and does not earn any points.

Fig. 1: Decision masks of a person I (left) and a person E1, E2 or E3 (right)

Your Decision for	Your payoff, if Person E chooses	
	X	Y
Option A	9	1
Option B	5	5

Your Decision for	Your payoff, if Person I chooses	
	A	B
Option X	9	5
Option Y	11	5

At the end of the round, the decision-maker E receives information about her payoff and person I's choice. Person I receives information about her payoff and in case she chose action A, she also learns E's decision (X or Y). In case person I chose B, she does not receive information about the E's decision. The passive assistants P1 and P2 do not receive any information until after the last round of the last interaction of the experiment. Then they are informed about their additional earnings.

All rounds are identical in terms of the procedure. The progress of the current interaction is displayed in tabular form in each round for the committee to which you are currently assigned.

End and final payoff

As soon as chance ends the last interaction, the experiment is over.

At the end of the experiment all interactions are paid off. The total amount of points from all rounds will be converted into Euros and paid out privately.

On the last screen of the last round of the last interaction, you can see how much you have earned in Euros.

Questions?

If you have any questions, please contact us. An experimenter will then come to your place.

If you think you have understood everything well, you may start the quiz on the screen. This quiz is only to make sure that everyone has understood the instructions well. The answers will not affect your payoff.

Additional Questions

Questions for the public:

PUB1. My decisions had a strong influence on the policy in subsequent periods.

- totally agree
- agree
- neutral
- do not agree
- totally do not agree
- no answer

PUB2. I had the impression that the fact that decision-makers were replaced by new individuals had a great influence on the outcome.

- totally agree
- agree
- neutral
- do not agree
- totally do not agree
- no answer

PUB3. How my choices would affect the policy-makers' payoffs was an important factor for my decisions.

- totally agree
- agree
- neutral
- do not agree
- totally do not agree
- no answer

PUB4. I made my choices in order to reward or punish the behavior of the policy-makers.

- totally agree
- agree
- neutral
- do not agree
- totally do not agree
- no answer

Questions for the committee members:

POL1. My individual choices had a strong influence on the behavior of the public in subsequent periods.

- totally agree
- agree
- neutral
- do not agree
- totally do not agree
- no answer

POL2. How my choices would affect the public's payoffs was an important factor for my decisions.

- totally agree
- agree
- neutral
- do not agree
- totally do not agree
- no answer

POL3. I made my choices in order to reward or punish the behavior of the public.

- totally agree
- agree
- neutral
- do not agree
- totally do not agree
- no answer

B.6 Results From the Questionnaire

Answers to the post-experiment questionnaire are summarized in Table B.6.1. Answers "totally agree", "agree", "neutral", "do not agree" and "totally do not agree" are coded from 1 to 5. "No answer" is removed from the analysis.

Table B.6.1: Summary of answers to questionnaire

	PUB1	PUB2	PUB3	PUB4	POL1	POL2	POL3
C	2.83	3.71	2.00	4.50	3.66	3.34	3.59
	(1.17)	(0.95)	-	(0.93)	(1.29)	(1.13)	(0.84)
CST	3.00	2.80	3.00	4.33	3.57	3.31	3.46
	(1.41)	(1.64)	(1.10)	(1.21)	(1.22)	(1.18)	(0.92)
CA	3.62	3.33	2.43	4.00	3.23	3.31	3.54
	(1.19)	(1.53)	(0.79)	(1.15)	(1.14)	(1.12)	(0.90)

Note: This tables summarizes the mean and standard deviations of answers to the post-experiment questionnaire. Answers "totally agree", "agree", "neutral", "do not agree" and "totally do not agree" are coded from 1 to 5. "No answer" is removed from the analysis.

Table B.6.2: Summary of answers to questionnaire

	df	N	F-value	p
PUB1	2	20	0.79	.469
PUB2	2	15	0.70	.516
PUB3	2	16	1.49	.262
PUB4	2	21	0.40	.675
POL1	2	85	0.91	.407
POL2	2	109	0.03	.966
POL3	2	123	0.20	.822

We perform ANOVA to test whether the answers are different across treatments. The results are summarized in Table B.6.2. We find no treatment difference in answers to all questions. The public from treatments C, CST and CA thinks that their decisions have a neutral effect on the policies of the committee in the subsequent period. They find that the fact that the decision-makers are replaced by new individuals have a neutral effect on the outcome. As for the incentives for their decision-making, the public agrees that the decision-makers payoff is an important factor. However, they do not make decisions to reward or punish the decision-makers. For the decision-makers in treatments C, CST and CA, they do not think that their choices have a strong influence on the choice of the public in the subsequent round. The payoff of the public is not an important factor for their decision-making, and neither do they make choices to reward or punish the public.

Appendix C

C.1 PPE of Correlated Grim-trigger Strategy

We construct the PPE δ-threshold of a grim-trigger strategy under Cor. Players start with C and never deviate to D as long as public signals match. To show that cooperation is optimal, we show that the gain from cooperation is at least as high as the gain from defection.

Denote by V_g the value in reward state where both players opt for C, and V_d the value in punishment state where both deviate to D. Denote further by $u(C)$ and $u(D)$ the expected payoffs from playing C or D. Cooperation is optimal if the following inequality holds

$$u(C) + \delta V_g + \geq u(D) + \delta \left(2\epsilon(1-\epsilon)V_g + (1 - 2\epsilon(1-\epsilon))V_d\right) \tag{C.1}$$

The biggest difference between correlated and non-correlated grim-trigger strategy is that if both cooperate, signals always match under Cor and players stick to cooperative path. We represent the non-expected stage-game cooperation, betrayal, sucker and defection payoffs by c, b, s, d. $u(C)$ and $u(D)$ is thus explicitly given by

$$u(C) = (1-\epsilon)c + \epsilon s$$
$$u(D) = (1-\epsilon)b + \epsilon d$$

The continuation value depends on own continuation strategy and has the following recursive form

$$V_g = (1-\epsilon)c + \epsilon s + \delta V_g$$
$$V_d = (1-\epsilon)d + \epsilon b + \delta V_d$$

Solving for V_g, V_d and plugging them and $u(C)$ and $u(D)$ back into (C.1) yields

$$(1-\epsilon)(c-b)+\epsilon(s-d)+\frac{\delta}{1-\delta}(1-2\epsilon+2\epsilon^2)\left((1-\epsilon)(c-d)+\epsilon(s-b)\right)\geq 0 \tag{C.2}$$

We can solve for δ and rewrite the expression by replacing c, b, s, d with expected payoffs and normalize them into a function of g:

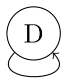
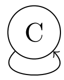
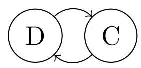
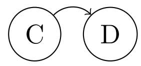
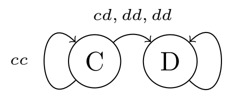
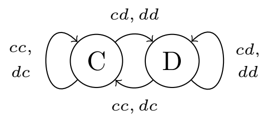
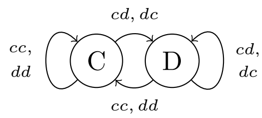
$$\delta \geq \frac{g}{1-2\epsilon+2\epsilon^2+g}.$$

Correlated grim-trigger strategy is a public perfect equilibrium if the above inequality holds.

C.2 Overview of Strategies

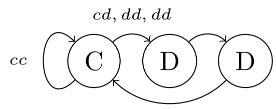
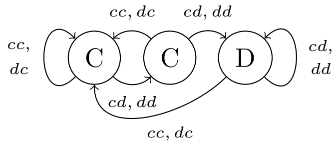
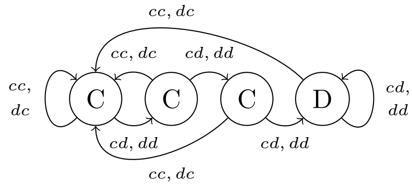
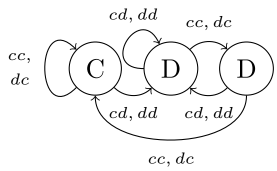
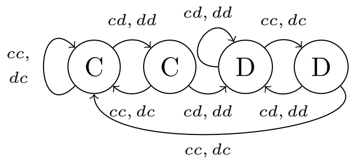
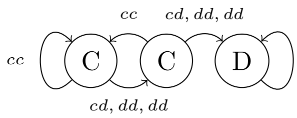
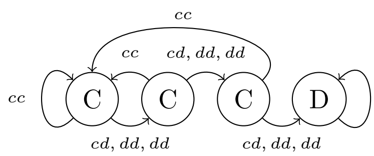
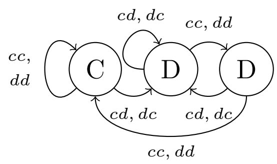
Tables C.2.1-C.2.3 list the 20 strategies taken from Fudenberg et al. (2012), and are reprinted from Appendix B of Dvorak and Fehrler (2018). The 20 strategies together with the four correlation-variational strategies CGRIM, CTFT, CWSLS and CT2 are used to produce Table C.3.1. Circles are strategy states and arrows indicate state transitions. Tables C.2.1-C.2.3 summarize the basic scenario of only five states. In our strategy frequency estimation, transition is assumed to condition on own action and both signals, thus nine states.

Table C.2.1: Strategies 1-7

Acronym	Description	Automaton
ALLD	Always play D.	
ALLC	Always play C.	
DC	Start with D, then alternate between C and D.	
FC	Play C in the first round, then D forever.	
Grim	Play C until either player plays D, then play D forever.	
TFT	Play C unless partner played D last round.	
PTFT (WSLS)	Play C if both players chose the same move last round, otherwise play D.	

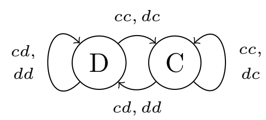
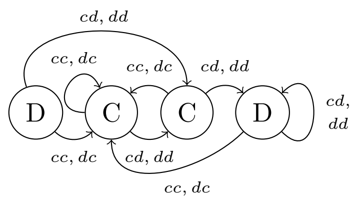
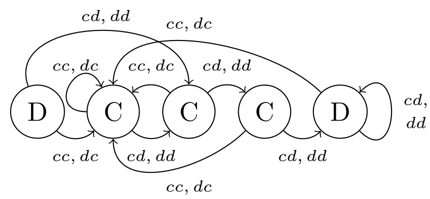
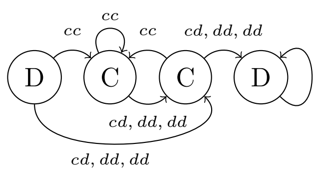
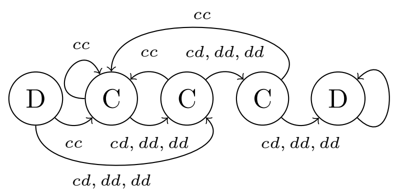
Notes: Circles represent the states of an automaton. The first state from the left is the start state. The labels C and D indicate whether the automaton prescribes cooperation or defection in the state. Arrows represent deterministic state transitions. The labels indicate the information profiles of the previous periods which trigger the transitions. An unlabeled arrow indicates an unconditional transition that occurs independent of the observed profile.

Table C.2.2: Strategies 8-15

Acronym	Description	Automaton
T2	Play C until either player plays D, then play D twice and return to C (regardless of all actions during the punishment rounds).	
TF2T	Play C unless partner played D in both of the last 2 rounds.	
TF3T	Play C unless partner played D in all of the last 3 rounds.	
T2FT	Play C unless partner played D in either of the last 2 rounds (2 rounds of punishment if partner plays D).	
T2F2T	Play C unless partner played 2 consecutive Ds in the last 3 rounds (2 rounds of punishment if partner plays D twice in a row).	
GRIM2	Play C until 2 consecutive rounds occur in which either player played D, then play D forever.	
GRIM3	Play C until 3 consecutive rounds occur in which either player played D, then play D forever.	
PT2FT	Play C if both players played C in the last 2 rounds, both players played D in the last 2 rounds, or both players played D 2 rounds ago and C last round. Otherwise play D.	

Notes: Circles represent the states of an automaton. The first state from the left is the start state. The labels *C* and *D* indicate whether the automaton prescribes cooperation or defection in the state. Arrows represent deterministic state transitions. The labels indicate the information profiles of the previous periods which trigger the transitions. An unlabeled arrow indicates an unconditional transition that occurs independent of the observed profile.

Table C.2.3: Suspicious Strategies 16-20

Acronym	Description	Automaton
DTFT	Play D in the first round, then play TFT.	
DTF2T	Play D in the first round, then play TF2T.	
DTF3T	Play D in the first round, then play TF3T.	
DGRIM2	Play D in the first round, then play GRIM2.	
DGRIM3	Play D in the first round, then play GRIM3.	

Notes: Circles represent the states of an automaton. The first state from the left is the start state. The labels C and D indicate whether the automaton prescribes cooperation or defection in the state. Arrows represent deterministic state transitions. The labels indicate the information profiles of the previous periods which trigger the transitions. An unlabeled arrow indicates an unconditional transition that occurs independent of the observed profile.

C.3 Additional Results

Table C.3.1: SFEM with Fudenberg et al. (2012) and Strategy Selection

	Cor	NoCor
ALLD	0.03	0.10
	(0.02)	(0.04)
ALLC	-	0.52
		(0.09)
DC	0.03	0.04
	(0.02)	(0.03)
FC	0.02	-
	(0.02)	
TFT	0.05	-
	(0.03)	
WSLS	0.04	-
	(0.03)	
TF2T	-	0.14
		(0.08)
GRIM2	0.31	0.16
	(0.09)	(0.07)
GRIM3	0.20	-
	(0.08)	
DTFT	0.02	-
	(0.02)	
CGRIM	0.21	0.04
	(0.08)	(0.03)
CTFT	0.11	-
	(0.06)	
γ	0.08	0.08
BIC	805.56	678.89
ln$L$	-381.39	-326.97

Notes: The table reports the maximum-likelihood shares of the strategies selected from Tables C.2.1-C.2.3 of Appendix B with data from the first 5 rounds of the last 3 supergames. All strategies condition on action-public signal profile $\{a_i, \omega_{-i}, \omega_i\}$. The estimation procedure assumes constant strategy use. γ is the estimated tremble probability, which avoids likelihood shares of zero when subjects deviate from a choice pattern. Strategies are selected based on Bayesian Information Criterion. Strategies attracting zero shares are omitted (-). Standard errors are reported in parentheses. Values may not add up to 1 because of rounding.

C.4 Communication Content

Table C.4.1: Frequency of Valid Coding for Sub-Categories

#	Subcategory	Category	Freq.	Freq. in Treatment Cor	NoCor	$\bar{\kappa}$
1	Proposal: both C	C	0.936	0.926	0.948	0.614
2	Proposal: both D	C	0.015	0.019	0.010	0.447
3	Proposal: alternate	C	0.078	0.111	0.042	0.756
4	Proposal: self D other C	C	0.054	0.074	0.031	0.872
5	Proposal: self C other D	C	0.025	0.046	-	0.762
6	Proposal: other coordination	C	0.025	0.037	0.010	0.479
7	Question: action of the other	C	0.005	0.009	-	0.164
8	Announcement: C	C	0.010	0.009	0.010	0.138
9	Announcement: D	C	0.005	0.009	-	0.328
10	Rejection of proposal	C	0.020	0.028	0.010	0.357
11	Acceptance proposal	C	0.711	0.639	0.792	0.368
12	Implicit punishment threat for D	C	-	-	-	-
13	Punishment threat grim	C	0.015	0.028	-	0.356
14	Punishment threat lenient grim	C	-	-	-	-
15	Approval of punishment threat	C	-	-	-	-
16	Ask for coordination	C	0.123	0.139	0.104	0.480
17	Benefits of C	D	0.142	0.157	0.125	0.310
18	Benefits of D	D	-	-	-	-
19	Benefits of asymmetric play	D	0.005	0.009	-	0.159
20	Related to fairness discussion	D	-	-	-	-
21	Related to strategic uncertainty	D	-	-	-	-
22	Related to payoffs	D	0.029	0.046	0.010	0.078
23	Related to Prisoner's dilemma	D	0.010	0.009	0.010	1.000
24	Related to game theory	D	0.005	0.009	-	1.000
25	Future benefit of C	D	-	-	-	-
26	Short term incentives of D	D	0.010	0.019	-	0.385
27	Related to signal comparison	D	0.005	0.009	-	0.187
28	Related to (un)matched signals	D	-	-	-	-
29	Promise	R	0.039	0.065	0.010	0.484
30	Distrust	R	0.025	0.046	-	0.289
31	Trust	R	0.005	-	0.010	0.235
32	Plea for trustworthy behavior	R	0.025	0.037	0.010	0.198
33	Implicit punishment threat when signals differ	C	-	-	-	-
34	Explicit punishment threat when signals differ grim	C	-	-	-	-
35	Explicit punishment threat when signals differ lenient grim	C	-	-	-	-
36	Small Talk	T	0.882	0.824	0.948	0.407
37	Off topic	T	-	-	-	-
38	Boredom	T	0.010	0.019	-	0.136
39	Confusion	D	0.010	0.019	-	0.035

Notes: Two raters identify whether a sub-category occurs in a subject-round observation. Frequency indicates the probability of occurrence of a sub-category whose classifications are the same between raters. The table shows overall frequency and frequency in treatments. Data is from last three supergames. The 39 sub-categories map into 4 main categories: Coordination (C), Deliberation (D), Relationship (R) and Trivia (T). Frequencies < 0.001 omitted (-). $\bar{\kappa}$ is the average Cohen's Kappa over all treatments. Mean $\bar{\kappa}$ of all subcategories with an overall frequency > 0.01 is 0.38.

C.5 Experimental Instructions and Quiz

[Below is the instructions and quiz for the information correlation treatment. Instructions for the other treatment is very similar and therefore is omitted. The original instructions are in German. Instructions for both treatments can be obtained from the authors upon request.]

Overview

Welcome to this experiment. We ask you not to talk to the other participants during this experiment and to switch off your mobile devices.

At the end of the experiment, You will be paid in cash for today's participation. The amount of money you receive depends on your own decisions, the other participants' decisions, and pure chance. It is important that you understand the instructions before the experiment starts.

In this experiment, every interaction between the participants runs through the computers you are sitting in front of. They will interact with each other anonymously. Neither your name nor the names of other participants will be announced. Also, for the evaluations only the anonymized data are used.

Today's session consists of several rounds. Your payout amount will be the sum of the points earned in all rounds, converted into euros. The conversion of the points into euros is done as follows. Each point is worth 2 cents, so that applies: 50 Points = 1.00 EUR.

All participants are paid privately, so other participants cannot see how much you have earned.

Experiment

Interactions and Role Assignment

This experiment consists of 7 interactions that are identical in their sequence, each consisting of a randomly determined number of rounds.

At the very beginning, before the first interaction, you will be randomly placed in a group with other participants. In each of the 7 interactions you will interact with another participant of your group.

Specifically, this is what happens: Before the first interaction, you will be assigned to a person from your group with whom you will interact in all rounds of the first interaction. In the second interaction, you will then be assigned to a new person from your group with whom you will interact in all rounds of the second interaction, and so on. In this way, you will interact with each person assigned to you in only one interaction, but in all rounds of that interaction.

Length of an Interaction

The length of an interaction is determined randomly. After each round there is an 80% chance that there will be at least one more payout-relevant round.

You can imagine this as follows. After each round, a 100-sided dice is thrown. If the roll results in a number less than or equal to 20, there is no further payout relevant round. If the roll is a different number (21-100), the interaction continues. Note that the probability of another payout-relevant round does not depend on the round you are in. The probability of a third payout-relevant round if you are in round 2 is 80%, as is the probability of a tenth payout-relevant round if you are in round 9.

A special feature concerns the first 5 rounds of each interaction. These rounds are always run even if the interaction has already been completed by the random number generator. At the end of the fifth round, you will find out whether the interaction has already been completed and, if so, up to which round your decisions were relevant for the payout. If the interaction has not been completed by round five, it will continue round after round and the interaction will be ended immediately if there is no further round.

When an interaction is finished, a new person is assigned for the next interaction. After the seventh interaction, the experiment ends.

Interaction and Round Schedule

At the beginning of an interaction, i.e. before the first round of the interaction, you can chat with the other person on the screen. The chat takes place in an anonymous chat window. To protect your anonymity, it is important that you do not give any information about yourself or your seat number while communicating. Otherwise we reserve the right not to pay you in the end. The chat content will be displayed during the interaction and you can read it.

Then the first round of interaction begins.

In each round you choose one of two possible options, A or B. The other person also chooses one of two possible options, A or B.

For each option, a signal is randomly determined in each round, which corresponds to the option with 80% probability. With 20% probability the signal does not correspond to the option but shows the other option. At the end of a round, you and the other person do not learn what the other person has chosen, but receive the signals determined for the chosen options. Your signal corresponds to the signal of the option chosen by the other person. The other person's signal corresponds to the signal of the option you have chosen. That is, if both persons choose the same option, both receive the same signal. If the two people have chosen different options, both can get different signals. This results in the following.

Important: Since exactly one signal is randomly determined for each option in each round, it may be possible to draw conclusions based on the two signals as to what the other person has chosen.

If two different signals occur, the other person has certainly chosen a different option than you. If you and the other person had chosen the same option, you and the other person would also receive the same signal (the signal determined for the option you both chose).

If two identical signals occur, the other person has either chosen the same option (case 1) or another option (case 2) and the two signals for the different options correspond randomly (probability for case 2: $0.8*0.2+0.2*0.8 = 0.32$).

Your income depends on the option you choose and the signal you receive. Similarly, the payout of the other person depends on the option you choose and the signal you receive.

Figure C.5.1: Round Income [Figure 1 from Instructions]

Ihre Optionen	Ihr Einkommen bei Signal		Erwartetes Einkommen, wenn die andere Person	
	A	B	Option A waehlt	Option B waehlt
Option A	32	2	26	8
Option B	40	10	34	16

In Figure 1, the four fields on the left indicate the lap income resulting from the combinations of the chosen option and your signal. The same table applies to the other person. For example, your lap income is 10 points if you chose option B and received signal B, and the other person's income is 2 points if he or she chose option A and received signal B.

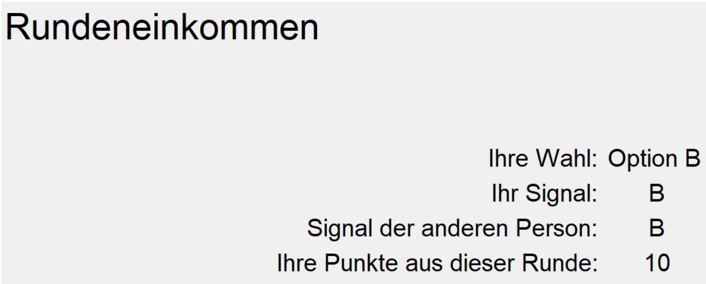
Once you and the other person have chosen an option, chance determines the signals for the options with the probabilities given above. The signals for the chosen options are then used to determine your lap earnings and those of the other person.

The four fields on the right in Figure 1 show the earnings you can expect depending on your option and the option of the other person. For example, if you choose option B and the other person chooses option A, you are 80% likely to receive signal A and 20% likely to receive signal B. Therefore, you will receive 40 points with 80% probability and 10 points with 20% probability, which means that your expected earnings in this case are: $0.8 * 40 + 0.2 * 10 = 34$ points.

At the end of the round you will receive a short feedback regarding your chosen option, the signal you received, the signal the other person received and your own round earnings (see Figure 2). You will not be informed of the other person's choice of option.

All possible subsequent rounds are identical in terms of the sequence of events. However, you can only chat with the other person before the first round of interaction. This is not possible before later rounds. The progress of the current interaction, i.e. the feedback that you received at the end of previous rounds, is displayed in a tabular view.

Figure C.5.2: Part of Feedback Screen (Example) [Figure 2 from Instructions]



End and Final Payoff

As soon as chance ends the last interaction, the experiment is over.

At the end of the experiment all interactions are paid off. The total amount of points from all rounds will be converted into Euros and paid out privately.

On the last screen of the last round of the last interaction, you can see how much you have earned in Euros.

Any Question?

If you have any questions, please contact us. An experimenter will then come to your place.

If you think you have understood everything well, you may start the quiz on the screen. This quiz is only to make sure that everyone has understood the instructions well. The answers to the quiz will not affect your payoff.

Quiz [on screen]

[After completing the quiz, correct answers will appear on the next screen.]

1. How many interactions are there?

   [1,7, it is by chance]

2. What is the minimum number of rounds in an interaction (payment-relevant or not)?

   [1, 3, 5]

3. What is the probability that there will be another payout-relevant round of interaction when you are in round six of an interaction?

   [20%, 80%, 100%]

4. What is the probability that the signal corresponds to the actual option?

   [20%, 80%, 100%]

5. Suppose you choose option A, the other person receives the signal A.
   (a) You receive the signal B, which option did the other person choose?

   [certain option A, certain option B, cannot be stated with certainty]

   (b) You also get the signal A, which option did the other person choose?

   [certain option A, certain option B, cannot be stated with certainty]

   (c) If the other person chose option B, how likely was it for you to receive signal A?

   [20%, 80%, 100%]

6. You choose option B. Suppose the other person also chooses option B.
   (a) With what probability will the other person receive signal B?

   [20%, 80%, 100%]

(b) You receive signal A, which signal does the other person receive?

[signal A, signal B, cannot be stated with certainty]

(c) What is your round income when you receive signal A?

[10, 34, 40]

(d) What is the expected round income of the other person?

[16, 34, 40]

C.6 Exchanging Private Information to Sustain Cooperation in Noisy, Indefinitely Repeated Interactions

[Below is our pre-analysis plan registered in January 2020 on the AEA RCT Registry.]

Experimental Design

We implement different variants of a Prisoners' Dilemma game with imperfect monitoring in a laboratory experiment. In every round, two players choose their actions $a_i \in \{C, D\}$ simultaneously. Payoffs depend on the player's own action a_i and the received signal about the other player's action ω_{-i}. Under public monitoring, subjects are informed about $(a_i, \omega_i, \omega_{-i})$ at the end of every round. Under private monitoring, subjects are informed about (a_i, ω_{-i}) at the end of every round. The continuation probability δ of the repeated game is 0.8. In the treatments without correlation, signals are drawn independently for each of the chosen actions. Signals are noisy and indicate the wrong action with probability $\epsilon = 0.2$. Therefore, if both players play C, the probability that the two signals differ is 0.32. In the treatments with correlation, both players receive the same signals if they choose the same action. The signals are correct, that is: they indicate cooperation (defection) when both choose cooperation (defection), with probability $1 - \epsilon = 0.8$. However, if their actions differ, signals are drawn independently, as in the treatments without correlation.

In all treatments, subjects engage in a pre-play communication-stage before the first round of every supergame. In this stage, subjects can communicate via a chat-box interface for 120 seconds.

Under private monitoring, two treatments have a reporting stage, which is implemented in the form of a structured communication stage after every round. In this stage, subjects can report the received signal from the current round to their partner (or misreport it).

In every session of every treatment, subjects are randomly divided into 3 matching groups, with 8 each. Subjects play 7 supergames with pre-determined lengths. At the beginning of every supergame, each subject is matched with a new partner from his/her matching group using perfect stranger matching, so that they do not play with the same partner for a second time. To keep the length of supergames constant across treatments, we generated 3 sequences of random numbers beforehand, and used them to determine the length L_i of each supergame.[1] To increase the number of observations per supergame, we adapt the block-random-termination method (Fréchette and Yuksel, 2017). Subjects play a block of 5 rounds at the beginning of every supergame. If the true length L_i is smaller or equal than 5, the supergame ends at the end of round 5 and only the first L_i rounds are payoff

[1] We used Stata to generate 3 sequences of uniformly distributed random numbers between 0 and 1 with seeds 3, 4, and 5 (we used seeds 1 and 2 in: Dvorak and Fehrler, 2018). Denote the 3 sequences as $\{r_n\}_i = \{r_1, r_2, ..., r_x\}_i$, where $i = 3, 4, 5$ indicates the seed underlying the sequence and $n \in \mathbb{N}$. The first supergame has x_1 rounds if $r_{x_1} \leq 0.2$ and for all $n < x_1$, $r_n > 0.2$. The second supergame has $x_2 - x_1$ rounds if $r_{x_2} \leq 0.2$ and for all $x_1 < n < x_2$, $r_n > 0.2$. And so forth. The resulting (lengths of the) sequences are SQ1 $(2, 8, 1, 5, 7, 1, 7)$, SQ2 $(4, 2, 2, 21, 4, 3, 5)$ and SQ3 $(2, 3, 1, 1, 4, 6, 6)$.

relevant. If L_i is larger than 5, the supergame continues until round L_i has been reached and all rounds are payoff relevant. Before the end of round 5, subjects are not informed about whether the supergame ends or not.

Subjects are required to answer control questions before the game starts. At the end of the experiment, subjects answer a short survey to elicit basic socio-economic characteristics, such as age and gender.

Experimental Parameters

Figure 1 shows a screenshot of the decision interface.

Figure C.6.1: Stage Game Parameters [Figure 1 from the pre-analysis plan]

Ihre Optionen	Ihr Einkommen bei Signal		Erwartetes Einkommen, wenn die andere Person	
	A	B	Option A waehlt	Option B waehlt
Option A	32	2	26	8
Option B	40	10	34	16

Notes: Screenshot from the experiment. Payoffs are in experimental currency units with an exchange rate of 50 ECU = 1 EUR.

The left two columns depict the stage-game payoff in experimental currency units. The payoff parameters do not vary across treatments. The last two columns show the expected stage-game payoffs and are calculated given a fixed error rate of 0.2 among all treatments. The parameters are chosen such that two conditions are satisfied:

1) Under imperfect private monitoring with signal correlation, there is a quasi-public perfect (truth-telling) equilibrium (QPPE) if reporting is allowed, in which both players play a "reporting grim-trigger" strategy. The reporting grim-trigger strategy prescribes the following behavior: Start with C and report your received signals truthfully, continue cooperating as long as both reports in the previous round are the same, otherwise defect for all subsequent rounds. The reporting mechanism translates the private monitoring into (quasi) public monitoring. An analogous cooperative PPE exists under imperfect public monitoring, in which subject play the same grim-trigger strategy as the one sketched above (but without reporting).

2) No cooperative (Q)PPE exists if there is either no correlation or there is correlation but it cannot be detected due to the absence of a reporting stage.

Treatments

We implement up to five different treatments. We begin with collecting data for two treatments with imperfect-public monitoring: one with correlation (T1) and one without (T2). In case we find a statistically significant treatment difference (see next section for details on the test), we continue with two private-monitoring treatments with correlated signals: one with reports (T3) and one without (T4). In case, we find a statistically significant treatment difference between T3 and T4, we continue with the final treatment T5, which is a private-monitoring treatment without correlation but with a reporting stage.

T1 Signals are public and independent.

T2 Signals are public and perfectly correlated if both actions are the same.

T3 Signals are private and perfectly correlated if both actions are the same. Participants can publicly report their private signal after each round.

T4 Signals are private and perfectly correlated if both actions are the same. Participants cannot report signals.

T5 Signals are private and independent. Participants can publicly report their private signal after each round.

Hypotheses Tests, Power, and Further Analyses

In a previous study of communication and cooperation in a noisy, indefinitely repeated Prisoner's Dilemma with uncorrelated signals, we saw high cooperation rates in the first rounds of the supergame with pre-play communication but then a steady and strong decline over the subsequent rounds (Dvorak and Fehrler, 2018). In a pretest session of treatment T1, we again saw high cooperation rates in the first round and a decline afterwards. However, the decline was much weaker.

Based on these observations and the existence (or absence) of cooperative (Q)PPEs in the different treatments, we formulate our main hypotheses:

H1a: *The cooperation rate will be higher in T1 than in T2.* [2]

H1b: *The cooperation rate will be higher in T3 than in T4.*

[2]The correct statement is "The cooperation rate will be higher in T2 than in T1".

H1c: *The cooperation rate will be higher in T3 than in T5.*

We test the corresponding H0s by comparing the cooperation rates in the first 5 rounds of the last 3 supergames between the treatments. We run one-sided t-tests, for which we average the cooperation rates within each matching group and then take these averages as our independent observations.

Our simulations (see next paragraph) suggest that we will have enough power ($> 80\%$) to detect effect sizes of 10 percentage points with 9 matching groups with 8 participants each per treatment.

Simulations for Assessing the Statistical Power

In the simulations, we iterate the following process 10,000 times for various effect sizes Δ:

1. Create a data set of 8 (subjects per matching group) * 9 (number of matching groups per treatment) * 2 (treatments) observations and an indicator variable for treatment 2.
2. Draw random numbers from the Bernoulli distribution with a success probability that starts at 1 in round 1 and then linearly declines to 0.9 in round 5 for treatment 1.[3]
3. Draw random numbers from the Bernoulli distribution with a success probability that starts at 1 in round one and then linearly declines to $0.9 - \Delta$ in round 5 for treatment 2.
4. Average the random draws from rounds 1-5 within each matching group. These are the simulated average cooperation rates.
5. Run a one-sided t-test on the matching-group averages. Return the $p-$value.

Finally, we compute the share of the $p-$values smaller than 0.05, which gives us the statistical power. We check the accuracy of the procedure by running it 10'000 times with the same success probabilities in both treatments, which results in a share of $p-$values smaller than 0.05 of 0.049, which is close to the 5% that we would expect for this scenario. The simulation results indicate that the power increases in Δ and is 80.1% with $\Delta = 0.1$. The power remains similar if we introduce matching-group-specific variation in the slopes of the declining cooperation probabilities.

[3] 0.9 was the cooperation rate we observed in round 5 of the pretest of the T1 treatment.

Further Analyses

In addition to testing our three hypotheses, we explore subjects' strategies across the treatments to better understand the aggregate findings. These analyses are explorative in nature and we, therefore, refrain from specifying further hypotheses. The questions we are interested to explore are the following:

- How are participants' choices in the private treatments with reports (T3 and T5) influenced by the reports of the previous period?

- Are the strategies used in the private treatment with reports (T3 and T5) similar to the strategies used in the the public treatment with correlated signals (T1 and T2)?

For the treatments T2 and T3, we are particularly interested to assess how many participants use a reporting grim-trigger strategy. It will further be interesting to compare the estimated strategies in T2 and T3 to the strategies in T1 and T5 for which the reporting grim-trigger strategy is theoretically not supported.

For the analysis of strategies, we build on the strategy frequency estimation method (SFEM) introduced by Dal Bó and Fréchette (2011), and use the R package stratEst (Dvorak, 2018), which was first used in Dvorak and Fehrler (2018). The SFEM is frequently used to obtain maximum-likelihood estimates of the shares of a candidate set of strategies in experimental data. However, the results of the SFEM are specific to this set and it is hard to know ex-ante which strategies should be included. To circumvent this problem, we will compute Maximum Likelihood estimates for an endogenously determined number of strategies where the structure of each strategy is the result of a model-selection process. Thus we will infer the strategies from the data rather than imposing a predefined set of strategies. The process will always start with a large number of such generic strategies, which will then be reduced step-by-step using the integrated-completed-likelihood criterion (ICL-BIC, Biernacki et al. (2000)). The ICL-BIC is an entropy-based selection criterion for mixture-models which has been used to estimate the dimensionality of the strategy space in other settings before (Breitmoser, 2015).

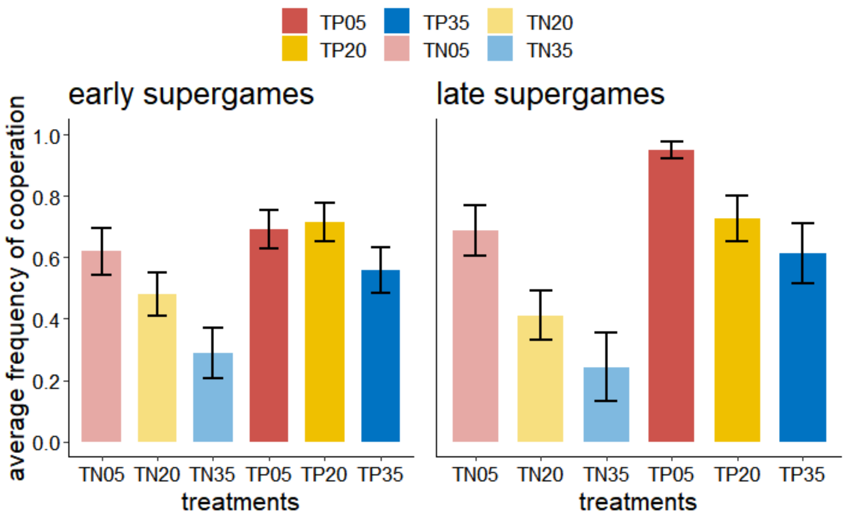
To assess whether strategies differ in two treatments, we fit a model on the pooled data of the two treatments and bootstrap the likelihood-ratio test statistic. If the distribution indicates that the likelihood-ration statistic is sufficiently extreme, we conclude that the strategies differ between the two treatments.

Appendix D

D.1 Cooperation Rates in Early and Late Supergames

One concern with adding an elicitation stage is that it may alter the subjects' behavior. In this section, I compare the average frequency of cooperation in supergames 2 and 3 with those in supergames 5 and 6, where elicitation tasks are introduced. Figure D.1.1 illustrates the average cooperation rates. Subjects are more cooperative in the late supergames in TP05 than in the early supergames ($p_{TP05} < .001$). They are as cooperative in the late supergames as in the early supergames in the other treatments ($p_{TP20} = .434$, $p_{TP35} = .294$, $p_{TN05} = .155$, $p_{TN20} = .173$, $p_{TN35} = .295$).

Figure D.1.1: Average rate of cooperation in the early and late supergames



Note: This figure shows the average cooperation rates in the early supergames (supergame 2 and 3) and late supergames (supergame 5 and 6). Data include the first five rounds in these supergames. Error bars represent standard deviations that are two-way clustered on both the subject ID and matching group ID.

D.2 Conditional Cooperation

Table D.2.1: Cooperation rates after memory-one histories of nine states

	$\sigma_{\emptyset}$	σ_{Ccc}	σ_{Ccd}	σ_{Cdc}	σ_{Cdd}	σ_{Dcc}	σ_{Dcd}	σ_{Ddc}	σ_{Ddd}	$\ln L$
TP05	0.83	0.96	0.96	0.64	0.50	0.25	0.41	0.00	0.24	-217.89
	(0.03)	(0.02)	(0.04)	(0.06)	(0.64)	(0.28)	(0.11)	(0.22)	(0.09)	
TP20	0.74	0.91	0.94	0.77	0.91	0.33	0.31	0.12	0.27	-285.02
	(0.06)	(0.03)	(0.04)	(0.06)	(0.05)	(0.11)	(0.07)	(0.09)	(0.08)	
TP35	0.66	0.91	0.86	0.75	0.67	0.26	0.30	0.14	0.12	-234.96
	(0.06)	(0.03)	(0.05)	(0.06)	(0.09)	(0.10)	(0.08)	(0.07)	(0.06)	
TN05	0.67	0.91	1.00	0.61	1.00	-	0.22	0.20	0.38	-241.54
	(0.08)	(0.04)	(0.13)	(0.07)	(0.69)		(0.06)	(0.27)	(0.10)	
TN20	0.53	0.79	0.75	0.46	0.47	0.24	0.12	0.16	0.22	-266.57
	(0.08)	(0.08)	(0.14)	(0.07)	(0.12)	(0.11)	(0.04)	(0.06)	(0.05)	
TN35	0.27	0.95	1.00	0.68	0.83	0.11	0.09	0.11	0.06	-115.59
	(0.09)	(0.07)	(0.05)	(0.10)	(0.20)	(0.09)	(0.04)	(0.05)	(0.03)	

Note: Estimated probabilities of cooperation after memory-one histories. The histories are action-signal combinations $\{a_i, \omega_j, \omega_i\}$. The numbers in the brackets are the standard errors. $\ln L$ is the log likelihood of the model.

D.3 Strategy Frequency Estimation With Behavior Strategies

Table D.3.1: Frequencies of pure and behavior strategies

	TP05	TP20	TP35	TN05	TN20	TN35
ALLC	-	-	-	0.14	-	-
				(0.10)		
ALLD	-	0.13	0.16	0.13	0.43	0.67
		(0.07)	(0.08)	(0.07)	(0.11)	(0.12)
GRIM2	-	0.34	-	-	-	-
		(0.10)				
TF2T	-	-	0.12	-	-	-
			(0.08)			
T2F2T	0.13	-	-	-	-	-
	(0.08)					
SGRIM	-	0.40	-	0.39	-	0.33
		(0.07)		(0.12)		(0.12)
M1	0.79	-	0.73	-	0.49	-
	(0.09)		(0.10)		(0.11)	
RAND	0.08	0.14	-	0.34	0.08	-
	(0.05)	(0.10)		(0.10)	(0.08)	
γ	0.06	0.09	0.15	0.03	0.14	0.07
BIC	400.52	511.17	481.54	430.86	498.85	220.98
ln$L$	-211.60	-246.92	-232.83	-207.49	-241.48	-106.29

Note: This table reports the maximum-likelihood shares of pure and behavior strategies. The estimation procedure assumes constant strategy use across all supergames. γ is the estimated tremble probability, which avoids likelihood shares of zero when subjects deviate from a choice pattern. Strategies are selected based on Bayesian Information Criterion. Strategies attracting zero shares are omiited (-). Standard errors are reported in parentheses. Values may not add up to 1 because of rounding.

This section reports the maximum-likelihood shares of the pure and behavior strategies. The three behavior strategies are SGRIM, M1BF and RAND. Subjects who play SGRIM play C if the action-signal pair was $\{C, c\}$ in the last round, and play D if the action-signal pair was $\{D, d\}$. They choose C with a probability $\sigma_{Cd} = \sigma_{Dc}$ if the action and signal differ. M1 is a memory-one strategy, similar to SGRIM. The only difference is that σ_{Cd} and σ_{Dc} are not necessarily identical. Rand is a randomization strategy with which the subject randomizes between C and D with $\sigma = 0.5$.

The estimated results are presented in Table D.3.1. For M1 in TP05, the estimated probabilities after history Cd and Dc are $\sigma_{Cd} = 0.70$ and $\sigma_{Dc} = 0.20$, which shows that the subjects are lenient but non-forgiving. For TP20, the estimated cooperation probability for SGRIM is $\sigma_{Cd} = \sigma_{Dc} = 0.95$. For M1 in TP35, the estimated probabilities after history Cd and Dc are $\sigma_{Cd} = 0.78$ and $\sigma_{Dc} = 0.27$. For treatments without pre-play communication, the cooperation probability is $\sigma_{Cd} = \sigma_{Dc} = 0.60$ for SGRIM in TN05. For M1 in TN20, the estimated probabilities after history Cd and Dc are $\sigma_{Cd} = 0.51$ and $\sigma_{Dc} = 0.21$. For TN35, the estimated probabilities are $\sigma_{Cd} = \sigma_{Dc} = 0.72$ for SGRIM.

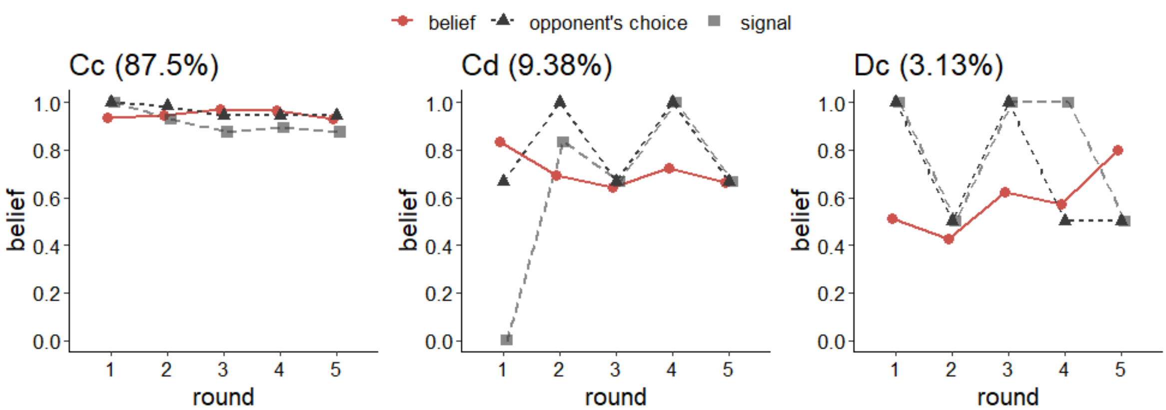
D.4 Conditional Beliefs

This section examines beliefs conditioning on round one action-signal pairs. I assume that the subjects condition their beliefs about their own actions and the signal representing the opponent's action. Figures D.4.1 to D.4.6 display the results.

At any noise level, communication raises beliefs in almost all cases. Beliefs are higher when the subject's own choice is C. This observation is not unique to treatments with pre-play communication. While subjects in the pre-play communication treatments can coordinate strategies such as always-cooperation or always-defection, subjects do not have the opportunity to coordinate their strategies if they do not engage in pre-play communication. However, subjects in the no communication treatments still hold high beliefs in the case Cd in which they themselves cooperate in round one but receive a defective signal. One possible explanation is that they tend to overestimate the probability of other players choosing the same action. This may also explain why the estimated proportion of ALLC is not low in TN35. According to Figure D.4.6, subjects have high beliefs following history Cd; however, they overestimate the likelihood that their opponent plays the same action.

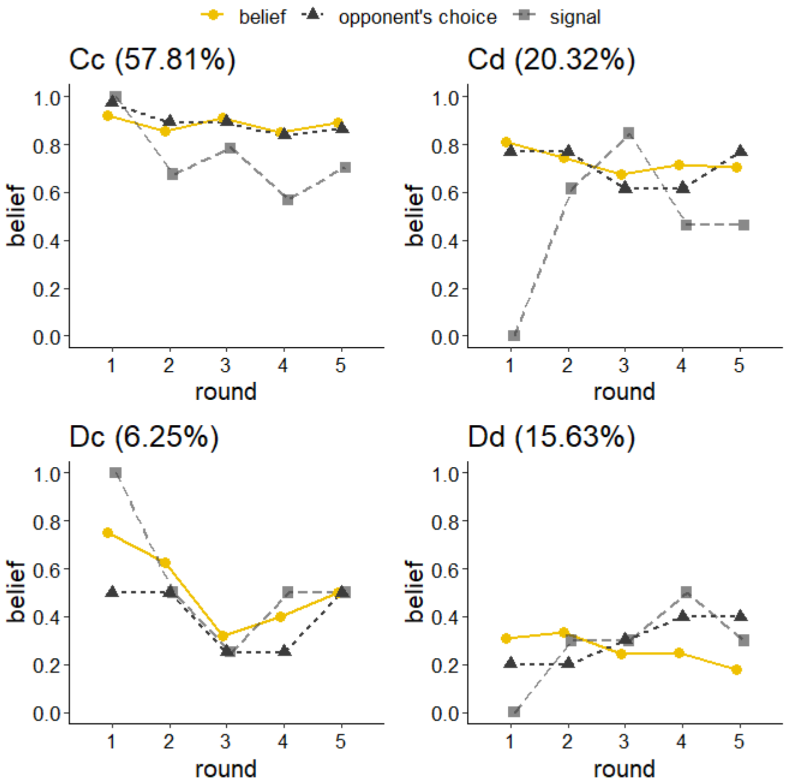
Subjects are not adjusting their beliefs to signals. The evolution of their beliefs is stabler than the evolution of the signals.

Figure D.4.1: Conditional beliefs for TP05



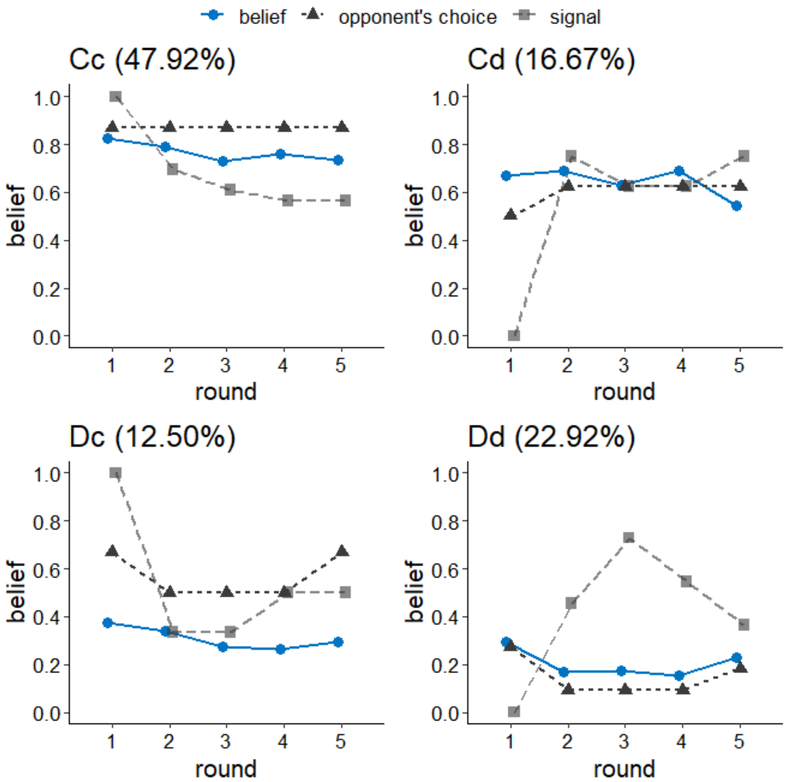
Note: Beliefs conditioning on round one histories. Histories are memory-one, which consist of one's own choice and the signal representing the opponent's choice. The evolution of opponents' signals and choices is also plotted in the Figures. Numbers in brackets are percentage of cases.

Figure D.4.2: Conditional beliefs for TP20



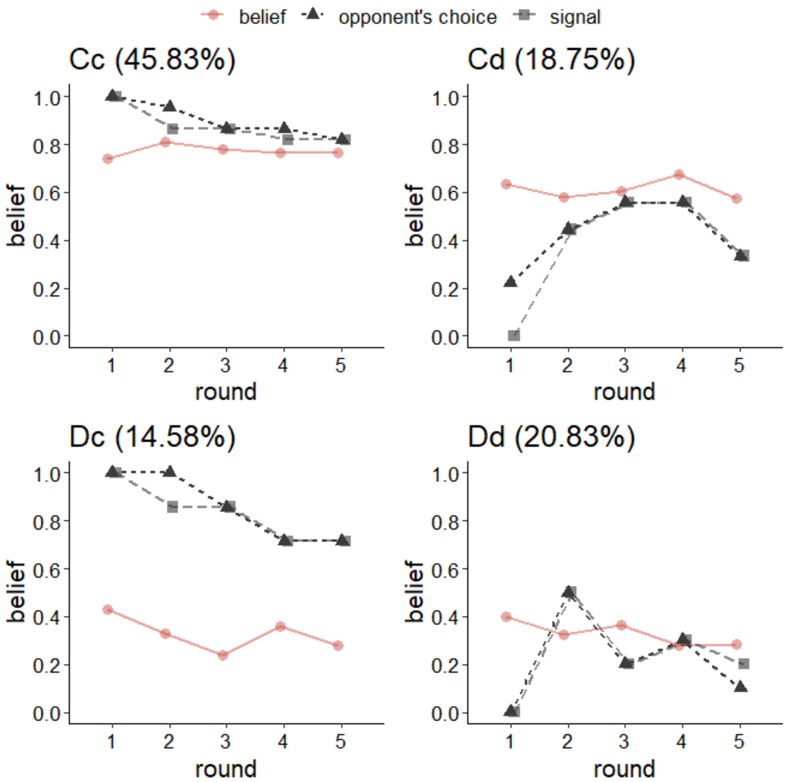
Note: Beliefs conditioning on round one histories. Histories are memory-one, which consist of one's own choice and the signal representing the opponent's choice. The evolution of opponents' signals and choices is also plotted in the Figures. Numbers in brackets are percentage of cases.

Figure D.4.3: Conditional beliefs for TP35



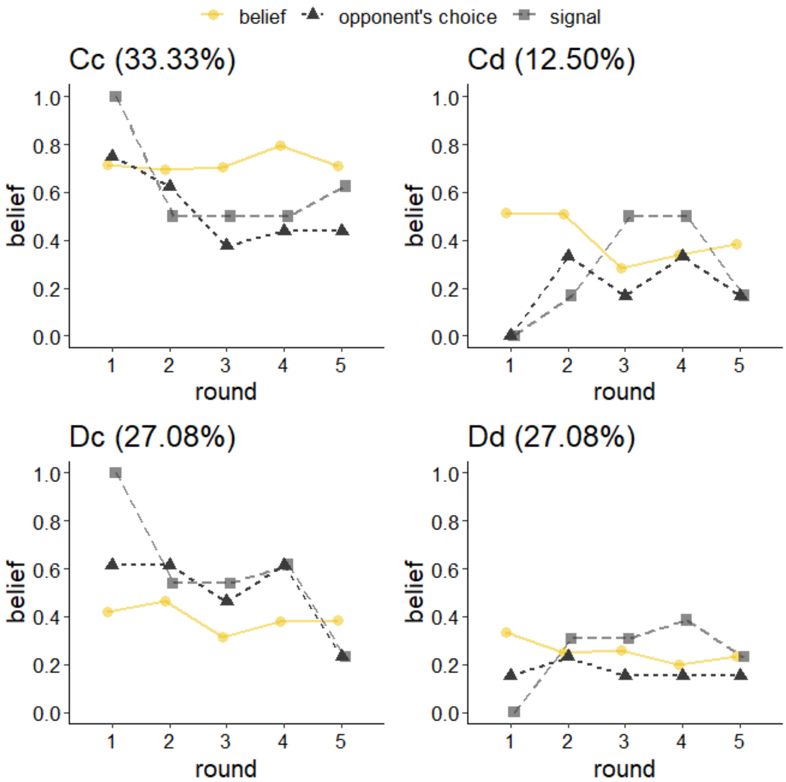
Note: Beliefs conditioning on round one histories. Histories are memory-one, which consist of one's own choice and the signal representing the opponent's choice. The evolution of opponents' signals and choices is also plotted in the Figures. Numbers in brackets are percentage of cases.

Figure D.4.4: Conditional beliefs for TN05



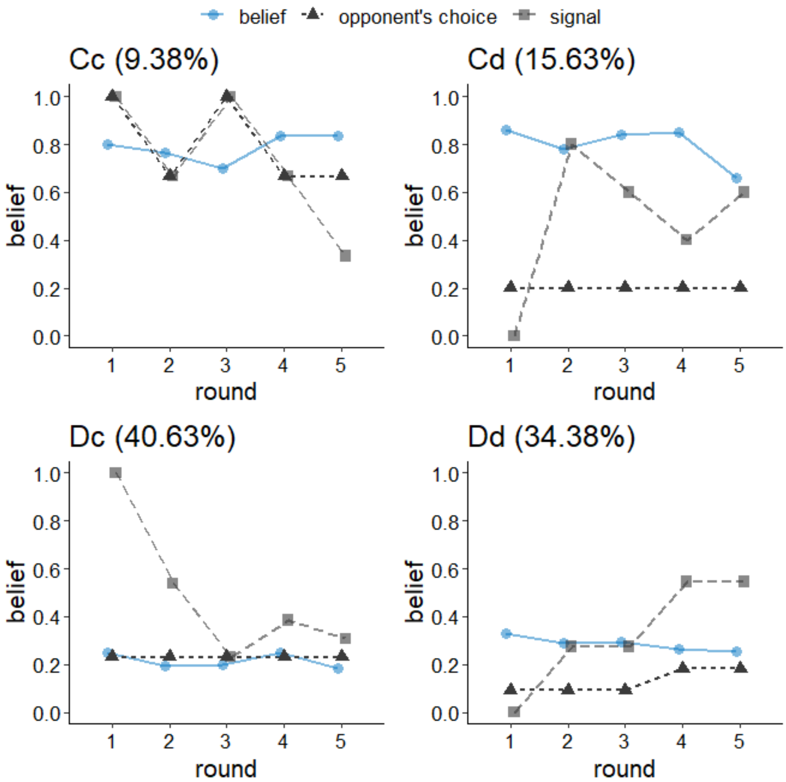
Note: Beliefs conditioning on round one histories. Histories are memory-one, which consist of one's own choice and the signal representing the opponent's choice. The evolution of opponents' signals and choices is also plotted in the Figures. Numbers in brackets are percentage of cases.

Figure D.4.5: Conditional beliefs for TN20



Note: Beliefs conditioning on round one histories. Histories are memory-one, which consist of one's own choice and the signal representing the opponent's choice. The evolution of opponents' signals and choices is also plotted in the Figures. Numbers in brackets are percentage of cases.

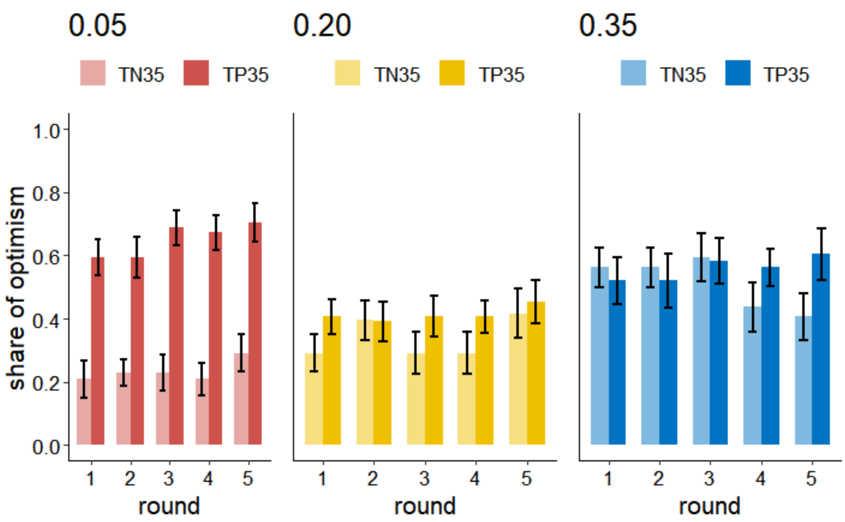
Figure D.4.6: Conditional beliefs for TN35



Note: Beliefs conditioning on round one histories. Histories are memory-one, which consist of one's own choice and the signal representing the opponent's choice. The evolution of opponents' signals and choices is also plotted in the Figures. Numbers in brackets are percentage of cases.

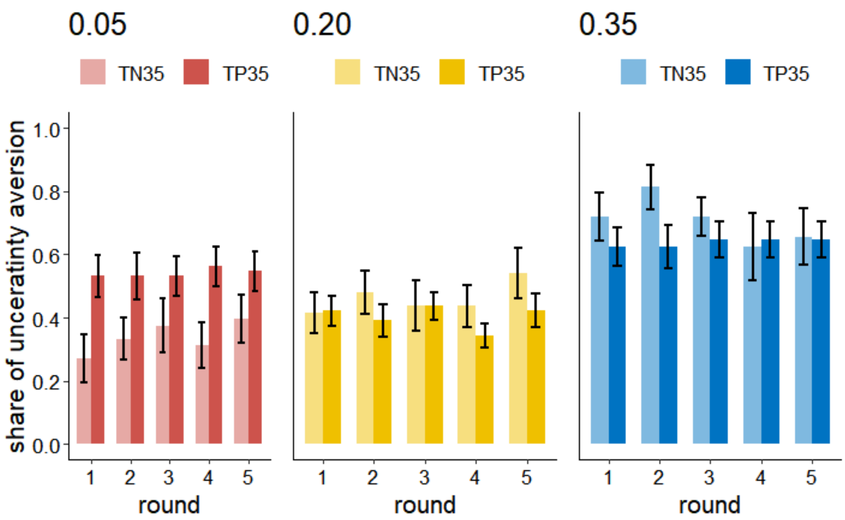
D.5 Attitudes Towards Strategic Uncertainty With Unrestricted Relative Risk Aversion

Figure D.5.1: Shares of optimistic subjects



Note: This figure shows the proportion of subjects who are optimistic about their beliefs. The samples are not restricted on the relative risk aversion parameter r.

Figure D.5.2: Shares of subjects who are ambiguity averse

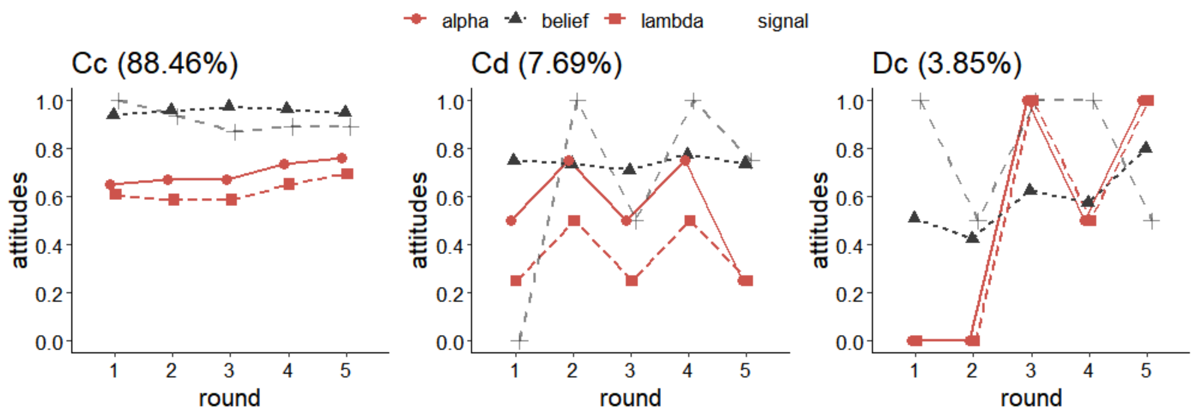


Note: This figure shows the proportion of subjects who are ambiguity averse. The samples are not restricted based on the relative risk aversion parameter r.

D.6 Conditional Attitudes

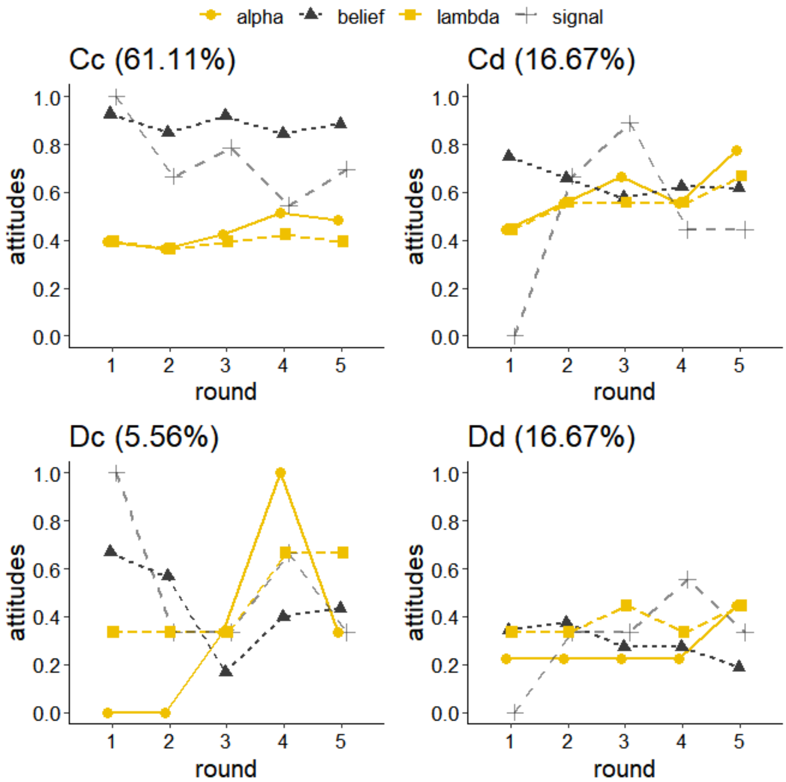
This section examines attitudes towards strategic uncertainty conditioning on round-one action-signal pairs. Subjects' attitudes towards strategic uncertainty consist of two parts: 1) α which measures how optimistic the subjects are regarding their beliefs, and 2) λ which measures the degree of ambiguity aversion. Figures D.6.1–D.6.4 show the results. The vertical axis of the figures shows the proportion of cases in which the subjects have positive α or positive λ. According to the figures, subjects' attitudes towards strategic uncertainty do not seem to relate to changes in beliefs and signals. They are neither more optimistic nor less ambiguity averse when they receive a cooperative signal in the initial round.

Figure D.6.1: Conditional attitudes for TP05



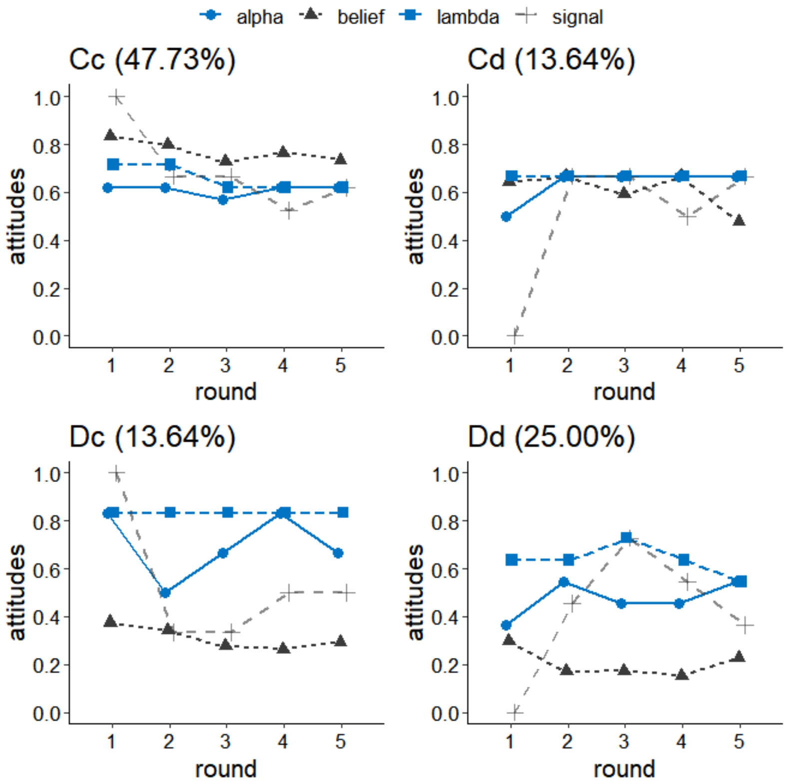
Note: The figures shows the proportion of positive α and λ conditioning on round one histories. Histories are memory-one, which consist of one's own choice and the signal representing the opponent's choice. The evolution of opponents' signals and beliefs is also plotted in the figures. Numbers in brackets are percentage of cases.

Figure D.6.2: Conditional attitudes for TP20



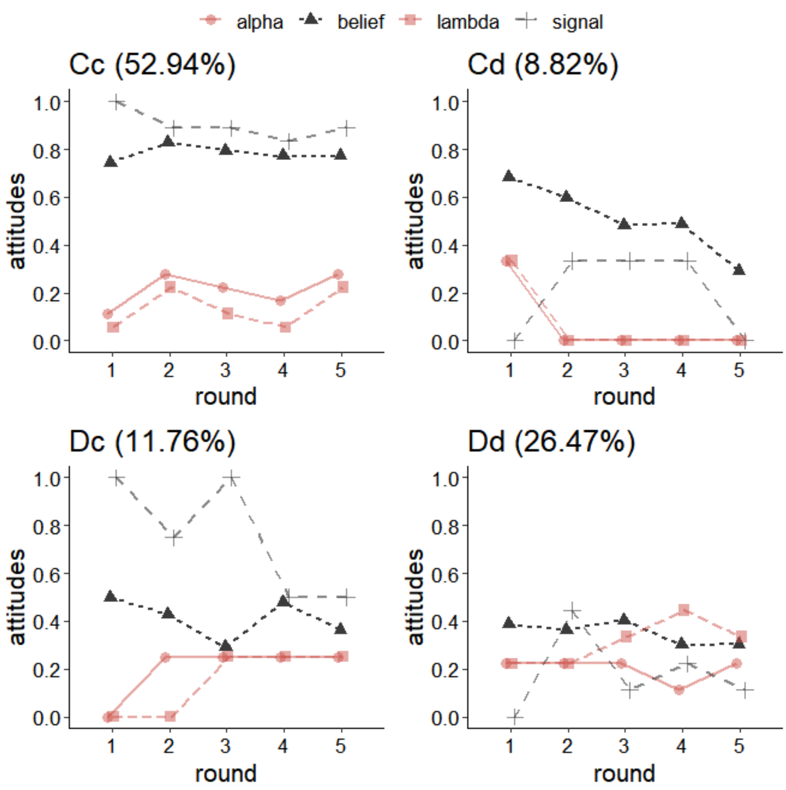
Note: The figures shows the proportion of positive α and λ conditioning on round one histories. Histories are memory-one, which consist of one's own choice and the signal representing the opponent's choice. The evolution of opponents' signals and beliefs is also plotted in the figures. Numbers in brackets are percentage of cases.

Figure D.6.3: Conditional attitudes for TP35



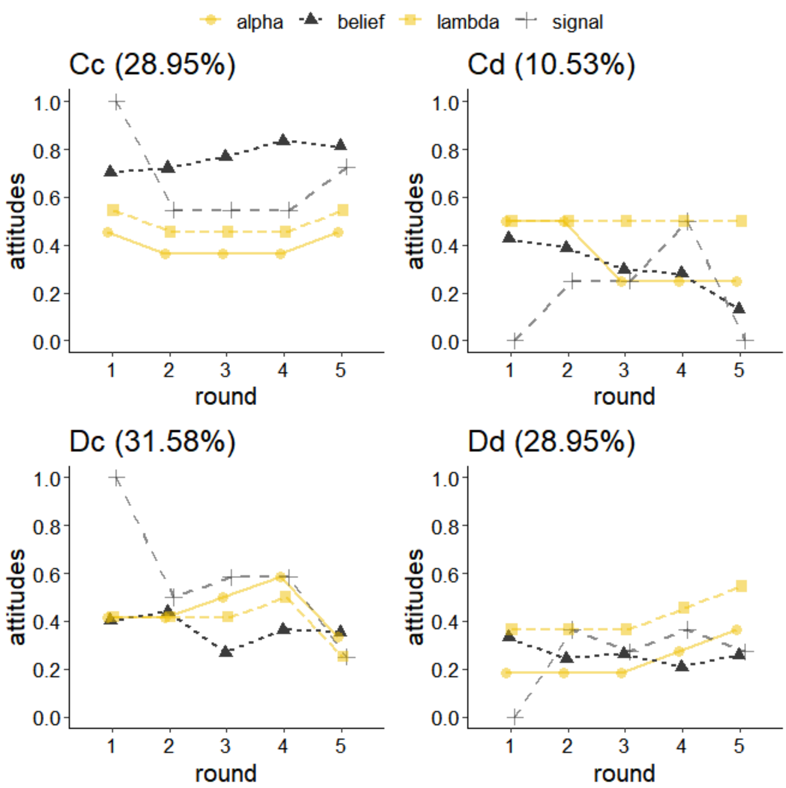
Note: The figures shows the proportion of positive α and λ conditioning on round one histories. Histories are memory-one, which consist of one's own choice and the signal representing the opponent's choice. The evolution of opponents' signals and beliefs is also plotted in the figures. Numbers in brackets are percentage of cases.

Figure D.6.4: Conditional attitudes for TN05



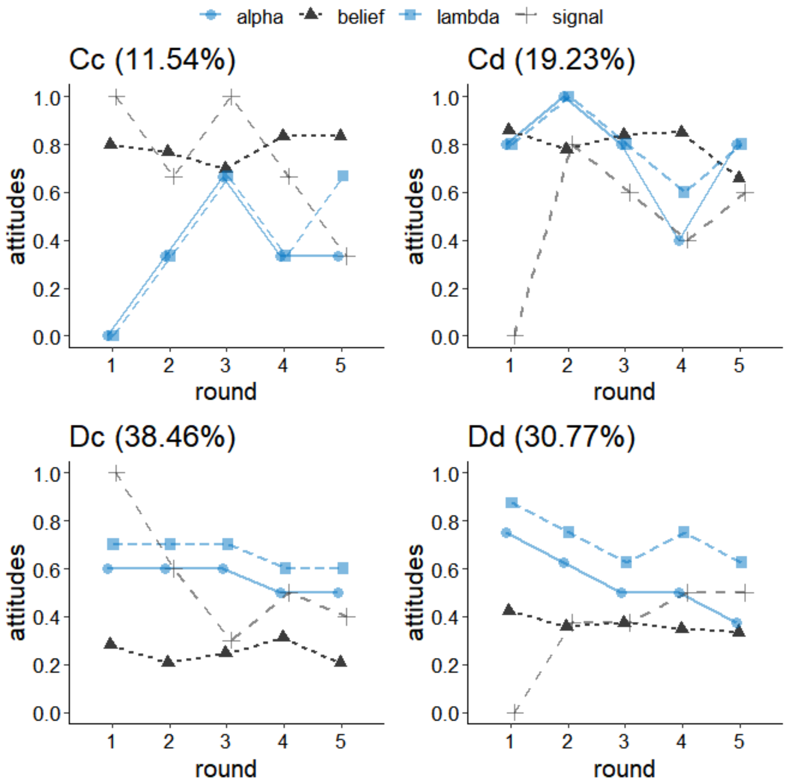
Note: The figures shows the proportion of positive α and λ conditioning on round one histories. Histories are memory-one, which consist of one's own choice and the signal representing the opponent's choice. The evolution of opponents' signals and beliefs is also plotted in the figures. Numbers in brackets are percentage of cases.

Figure D.6.5: Conditional attitudes for TN20



Note: The figures shows the proportion of positive α and λ conditioning on round one histories. Histories are memory-one, which consist of one's own choice and the signal representing the opponent's choice. The evolution of opponents' signals and beliefs is also plotted in the figures. Numbers in brackets are percentage of cases.

Figure D.6.6: Conditional attitudes for TN35



Note: The figures shows the proportion of positive α and λ conditioning on round one histories. Histories are memory-one, which consist of one's own choice and the signal representing the opponent's choice. The evolution of opponents' signals and beliefs is also plotted in the figures. Numbers in brackets are percentage of cases.

[Instructions and control questions for treatment TP05]

Overview

Welcome to this experiment. We ask you not to talk to other participants during the experiment and to turn off your mobile devices.

For today's participation you will be paid in cash at the end of the experiment. The amount of the payout is partly dependent on your decisions as well as the decisions of the other participants, the rest is a matter of luck. It is therefore important that you understand these instructions before starting the experiment.

In this experiment, participants interact anonymously via the interacting computers in front of you. Neither your name nor the names of other participants are disclosed. Also only the anonymized data will be used for the evaluations.

Today's experiment consists of several rounds. After the experiment, you will receive the sum of points earned (in euros). The conversion of points into euros is done as follows. Each point is worth 4 **cents**, i.e.:

50 points = 2 EUR.

Payment is made confidentially by bank transfer, so the other participants cannot see how much you have earned.

Experiment

Interactions and allocation

This experiment includes 6 interactions, each consisting of a randomly determined number of rounds. In the interactions, certain tasks are to be performed, which are explained in the following.

Right at the beginning, before the first interaction, you will be randomly assigned with other participants. In each of the 6 interactions you will interact with another participant.

Procedure: Before the first interaction, you are assigned to a person with whom you interact in all rounds of the first interaction. Then, in the second interaction, you are assigned to a new person with whom you interact in all rounds of the second interaction.

Duration of an interaction

The duration of an interaction is determined randomly. After each round, there is at least one more payout-relevant round with 80% probability.

You can imagine this as follows: After each round, a 100-sided die is rolled. If a number less than or equal to 20 is rolled, there is no further payout round. If any other number (21-100) is rolled, the interaction continues. Note that the probability of another payout-relevant round does not depend on which round you are in. The probability of a third payoff-related round if you are in round 2 is 80%, as is the probability of a tenth payoff-related round if you are in round 9.

It should also be noted that the first 5 rounds of each interaction are always carried out, even if the interaction has already been ended by the random number generator. At the end of the fifth round you will then find out whether the interaction has already ended and if so, up to which round your decisions were relevant for payout. If the interaction has not ended by round five, it continues round by round. The interaction ends when there is no more round.

As soon as an interaction is finished, a new person is assigned to you for the next interaction. After the sixth interaction, you perform a lottery task alone - then the experiment ends.

Tasks during an interaction

Before the first round of each interaction, you can chat with the other person on the screen. The chat takes place in an anonymous chat window. To protect your anonymity, it is important that you do not share any information about yourself or your seat number during the communication. Otherwise, we reserve the right not to pay you money in the end.

<u>Your task for interaction 1 to 3</u>

In each round, you choose one of the options A or B. At the same time, the other person also chooses one of the options A or B.

For each selection, a signal is randomly determined in each round that corresponds to the selected option with 95% probability. With 5% probability, the signal does not correspond to the actually selected option, but shows the other option. Neither you nor the other person will know at the end of a round what the other person has chosen - but you will receive the determined signals. Your signal corresponds to the signal determined for the other person's choice. The other person's signal corresponds to the signal determined for your selection. This means that if the other person chose option A, for example, you will receive a correct signal A with 95% probability, and you will receive an incorrect signal B with 5% probability. If you chose option A, the other person will receive the correct signal A with 95% probability and signal B with 5% probability.

Your round income depends on the option you have chosen and the signal you have received. Likewise, the other person's round income depends on his chosen option and the signal he received.

In Figure 1, the four boxes on the left (headed “Round income”) show the round incomes resulting from the combinations of Option and Signal. The same table applies to the other person. For example, your round income is equal to 12 points if you chose option B and received signal B, and the other person's income is equal to 4 points if he chose option A and received signal B.

In the four fields on the right in figure 1 you will find the income to be expected for you depending on your option choice and the option choice of the other person. For example, if you choose option B and the other person chooses option A, you will receive signal A with 95% probability and signal B with 5% probability. Therefore, with 95% probability you will receive 38 points and with 5% probability you will receive 12 points, i.e. your expected earning in this case is: 0.95*38+0.05*12=36.7 points.

Fig.1 Round income

Ihre Optionen	Ihr Einkommen bei Signal		Erwartetes Einkommen, wenn die andere Person	
	A	B	Option A waehlt	Option B waehlt
Option A	32	4	30.60	5.40
Option B	38	12	36.70	13.30

At the end of the round, you will receive brief feedback on the option you chose, the signal you received, and the signal the other person received; you will also receive feedback on your round income (see Figure 2). You will not be told about the other person's option choice.

Fig.2 Feedback screen section (example)

Rundeneinkommen

Ihre Wahl:	Option A
Ihr Signal:	A
Signal der anderen Person:	A
Ihre Punkte aus dieser Runde:	32

All subsequent rounds are identical in terms of their sequence. The course of the current interaction that you have received at the end of each of the previous rounds is tabulated in each round.

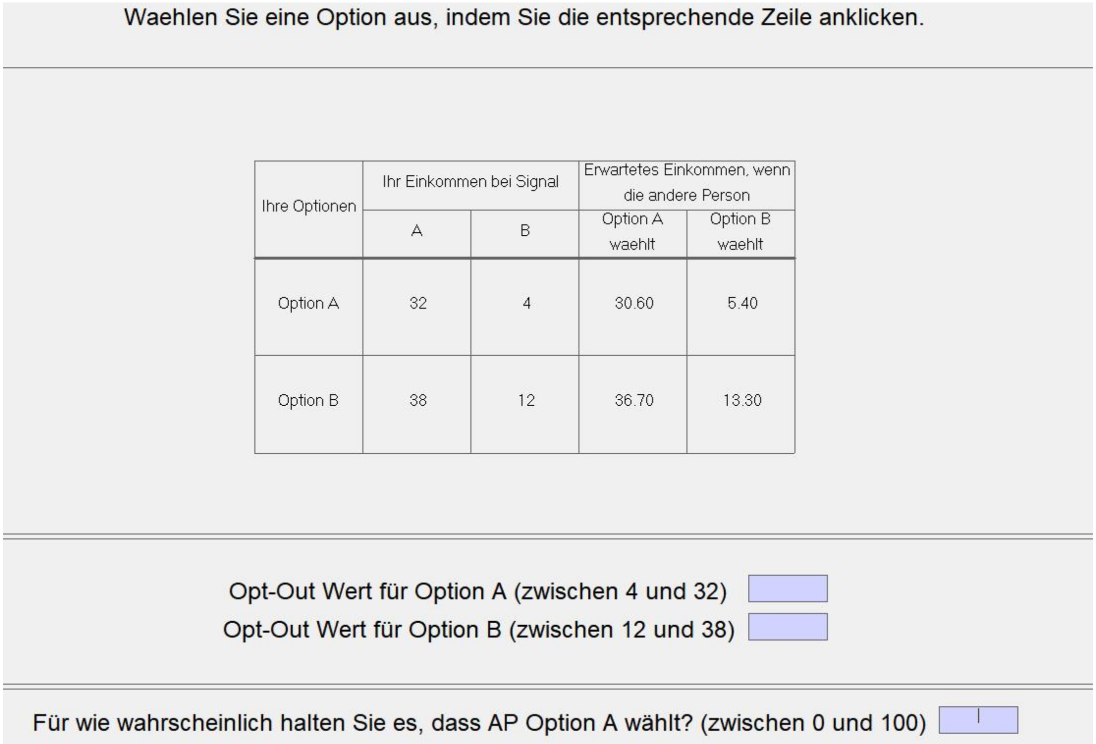
<u>Payment Interactions 1 to 3</u>

In Interactions 1 to 3, all the points you earned in the payout-relevant rounds count towards the payout at the end of the experiment.

<u>Your tasks for interaction 4 to 6</u>

You have four tasks in each round of interactions 4 through 6. Figure 3 is a screen shot of all 4 tasks.

Fig.3 Tasks in interactions 4 to 6



Waehlen Sie eine Option aus, indem Sie die entsprechende Zeile anklicken.

Ihre Optionen	Ihr Einkommen bei Signal		Erwartetes Einkommen, wenn die andere Person	
	A	B	Option A waehlt	Option B waehlt
Option A	32	4	30.60	5.40
Option B	38	12	36.70	13.30

Opt-Out Wert für Option A (zwischen 4 und 32)

Opt-Out Wert für Option B (zwischen 12 und 38)

Für wie wahrscheinlich halten Sie es, dass AP Option A wählt? (zwischen 0 und 100)

Task 1: It is the same as in interactions 1 to 3. You choose between option A and option B, and your round income is determined by the option you choose and the signal you receive.

Task 2: You specify an opt-out value for Option A by answering the following question:

Suppose the computer decides for you in this round and chooses option A. How many points would you like to have at least in order not to (have to) proceed with option A?

You enter an amount between 4 and 32 points with up to two decimal places. The computer randomly selects an amount between 4 and 32 points with up to two decimal places. Your income is determined by this randomly selected amount as follows:

If the randomly selected amount is greater than or equal to the opt-out value you specify, your income will be equal to the amount randomly selected by the computer.

If the amount chosen at random is less than your specified opt-out value, your income will be determined by option A from the computer and the signal from the other person's option from task 1.

Example: you have specified 27.5 and the randomly chosen amount is 30.76. Since the randomly chosen amount is larger than your opt-out value, your income is 30.76. However, if the randomly chosen amount is 10.24, which is smaller than your opt-out value, and suppose you received signal A in this round, then your income is determined by option A and signal A, which is 32 points (see Figure 1).

Task 3: You specify an opt-out value for option B by answering the following question.

Suppose the computer decides for you in this round and chooses option B. How many points would you like to have at least in order not to (have to) proceed with option B?

You enter an amount between 12 and 38 points with up to two decimal places. The computer randomly selects an amount between 12 and 38 points with up to two decimal places. Your income is determined by this randomly selected amount.

If the randomly selected amount is greater than or equal to the opt-out value you specify, your income will be equal to the amount randomly selected by the computer.

If the amount chosen at random is less than your specified opt-out value, your income will be determined by option B and the signal of the other person's option from task 1.

Example: You have specified 20.56 and the randomly selected amount is 20.56. Since the randomly selected amount is equal to your opt-out value, your income is 20.56 points. However, if the randomly chosen amount is 19.5, which is less than your opt-out value, and suppose you received signal A in this round, then your income is determined by option B and signal A, which is 38 points (see Figure 1).

Task 4: You answer the question with what probability you think the other person chose option A in this round. The estimate given should be an integer between 0 and 100. Your estimate will decide whether you win 50 points or not. The rules that decide whether you win have been set up in such a way that you have very high chances of winning if you are really convinced of their answer. If you want to know the exact mechanism, please ask the experimenter after the experiment.

<u>Payment Interactions 4 to 6</u>

Your payment in interactions 4 to 6 is determined as follows.

One interaction is selected at random. The amount of the payout in this interaction is determined by **task 1** (see above), i.e. by the income of all rounds relevant for the payout.

Another interaction is randomly selected by the computer, and your payout in that interaction is determined by **Task 4.** However, not all rounds are payouts, the computer randomly selects one round and you are only paid in that round.

Your payoff in the remaining interaction is determined by either Task 2 or Task 3. The computer randomly chooses a round and then between Task 2 and Task 3 in that round. Your payoff in this interaction is determined by either Task 2 or Task 3 in the selected round.

During the processing of the tasks you do not know which interaction is paid by which task. At the end of each round, you will only receive feedback as in Figure 4. You will only receive detailed information about your payoff at the end of the experiment.

Fig.4 Feedback in interactions 4 to 6

Lottery task

When all 6 interactions are finished, you are faced with 22 lottery tasks. 11 of them pay out either 4 or 32 points, the others either 12 or 38. For all lottery tasks, the probability of getting the higher of the two possible winnings ranges from 0% to 100% in increments of 10%. See Figure 5.

Fig.5 Lottery task

Die Wahrscheinlichkeit, mit der der Computer die höhere Auszahlung wählt	Opt-Out Wert für die Lotterie, die entweder 4 oder 32 Punkte auszahlt	Opt-Out Wert für die Lotterie, die entweder 12 oder 38 Punkte auszahlt
0%		
10%		
20%		
30%		
40%		
50%		
60%		
70%		
80%		
90%		
100%		

For each of the 22 lottery tasks, enter an opt-out value by answering the question:

If you were allowed to choose between getting a certain number of points safely or playing the lottery; what is the minimum number of points you would like to have in order to decide against playing the lottery?

Your opt-out value should be between 4 and 32 for lottery tasks that pay out either 4 or 32, and between 12 and 38 for lottery tasks that pay out either 12 or 38. Your opt-out value can be up to two decimal places.

Your payment in the lottery task is determined in the following two-step process. First, the computer randomly selects one of the 22 lottery tasks. Second, the computer randomly selects an amount between 4 and 32 if that selected lottery pays either 4 or 32, or randomly selects an amount between 12 and 38 if that selected lottery task pays either 12 or 38.

If this randomly chosen amount is greater than or equal to your opt-out value, you will receive this random amount as a payout.

If this randomly chosen amount is less than your opt-out value, your payment will be determined by the lottery.

Example: The computer chooses the lottery task that pays out either 12 or 38 points, and the probability of getting 38 points is 70%. Let's assume that your specified opt-out value for this lottery is 30.55:

If the random amount drawn by the computer is less than 30.55, for example 20.71, then your payout will be determined by the lottery. That is, there is a 70% chance that you will receive 38 points and a 30% chance that you will receive 12 points.

If the random amount drawn by the computer is greater than or equal to your opt-out value, for example 32.07, then you will receive those 32.07 points.

End and payment

When the lottery task is finished, the experiment is over.

Your total payout is composed of 5 parts. The first part consists of all the points you earned in the payout relevant rounds in interactions 1 to 3. The second part consists of all the points you earned in the payout relevant rounds in a randomly selected interaction in interactions 4 to 6. The third part consists of the points you earned in a randomly selected round in an interaction, also randomly selected, in interactions 4 to 6 from task 4. The fourth part consists of the points you earned from task 2 or task 3 in a round in a randomly selected interaction from 4 to 6. In addition, there are the points from the lottery.

At the end of the experiment, the points income is converted into euros. The payment will be made confidentially by bank transfer. After the lottery task, you will see on the screen how much you have earned in euros.

Questions?

Take the time to go through the instructions again thoroughly. If you have any questions, please let us know. An experiment leader will then come to your place.

If you have understood everything well, you can now start the quiz on the screen. The quiz is just to make sure that you have understood the instructions well. The answers of the quiz will not affect your payment.

Quiz [on screen]

[After completing the quiz, correct answers will appear on the next screen]

1. How many interactions are there?

[3, 6, it is by chance]

2. What is the minimum number of rounds in an interaction (payment-relevant or not)?

[1, 3, 5]

3. What is the probability that there will be another payout-relevant round of interaction when you are in round six of an interaction?

[20%, 80%, 100%]

4. What is the probability that the signal corresponds to the actual option?

[20%, 80%, 100%]

5. Suppose you choose option A, the other person receives the signal A.

(a) You receive the signal B, which option did the other person choose?

[certainly option A, certainly option B, cannot be stated with certainty]

(b) Suppose the other person choose option B. What is the probability for you to receive signal A?

[5%, 20%, 95%]

6. Suppose you are in interaction 4. You choose option B and receive signal A. Suppose the other person also chooses option B.

(a) How many tasks are there in interaction 4?

[1, 3, 4]

(b) What signal does the other person receive?

[signal A, signal B, cannot be stated with certainty]

(c) What is your round income from task 1 when you receive signal A?

[4, 32, 38]

(d) Suppose your opt-out-value for task 2 is 30. The random number drawn by the computer is 31.52. What is your income from task 2?

[31.52, 30, 32]

(e) Suppose your opt-out-value for task 2 is 30. The random number drawn by the computer is 29.22. What is your income from task 2?

[32, 30, 29.22]

(f) During the interactions, do you know which interaction will be paid by which task?

[Yes, No]

7. Is there another task after the six interactions?

[No, Yes a lottery task]

D.7 Questionnaire

The following questions are asked in the second part of the questionnaire in TP05. Subjects answer with slider bars of scale 0-1, where 0 stands for "totally do not agree". The first part of the questionnare consists of standard socio-demographic questions.

1. (Impact) What I and the other person have agreed on during communication has a big impact on my decisions.
2. (Accordance) I make decisions in accordance to what I and the other person have agreed on during communication.
3. (Reward) I make decisions in order to reward the other person.
4. (Punish) I make decisions in order to punish the other person.
5. (More point) I make decisions in order to help the other person earn more points.
6. (Less point) I make decisions in order to stop the other person earn more points.
7. (Long-run) I make decisions in order to earn more points in the long run.
8. (Close) In general, I feel close to the the other players.
9. (Trust) I trust the other players.
10. (Interpret) Imagine that the other person has agreed with you during communication to play Option A. Suppose in one round, your signal was B. How likely do you think this signal is due to error instead that the other player has intentionally chosen Option B?

D.8 Answers to Questionnaire

Answers to the post-experiment questionnaire are summarized in Table D.8.1. The first six columns summarize the means and standard deviations of the answers for each treatment. The ANOVA test results are shown in the last four columns. Significant treatment differences are detected in questions regarding rewards, trust, and the interpretation of a defective signal.

Table D.8.1: Summary of answers to questionnaire

	TP05	TP20	TP35	TN05	TN20	TN35	df	N	F-value	p
Impact	0.84	0.88	0.84	-	-	-	2	88	0.20	.821
	(0.32)	(0.23)	(0.27)							
Accordance	0.85	0.81	0.82	-	-	-	2	88	0.19	.827
	(0.28)	(0.30)	(0.29)							
Reward	0.55	0.48	0.45	0.40	0.30	0.15	5	152	3.44	.006
	(0.38)	(0.41)	(0.36)	(0.38)	(0.26)	(0.24)				
Punish	0.11	0.21	0.11	0.19	0.28	0.19	5	152	1.33	.255
	(0.23)	(0.33)	(0.20)	(0.29)	(0.33)	(0.28)				
More point	0.44	0.42	0.49	0.31	0.28	0.22	5	152	2.06	.074
	(0.33)	(0.40)	(0.34)	(0.36)	(0.31)	(0.28)				
Less Point	0.10	0.12	0.11	0.09	0.17	0.13	5	152	0.40	.849
	(0.20)	(0.25)	(0.19)	(0.15)	(0.26)	(0.24)				
Long-run	0.79	0.74	0.74	0.73	0.67	0.60	5	152	0.85	.517
	(0.34)	(0.32)	(0.31)	(0.32)	(0.32)	(0.39)				
Close	0.43	0.39	0.32	0.33	0.25	0.18	5	152	1.98	.085
	(0.32)	(0.34)	(0.25)	(0.35)	(0.32)	(0.24)				
Trust	0.73	0.60	0.55	0.46	0.29	0.20	5	152	9.39	$< .001$
	(0.31)	(0.35)	(0.30)	(0.35)	(0.29)	(0.23)				
Interpret	0.62	0.54	0.58	0.26	0.33	0.38	5	152	5.11	$< .001$
	(0.36)	(0.36)	(0.31)	(0.32)	(0.29)	(0.32)				

Note: This table summarizes the means and standard deviations of the answers to the post-experiment questionnaire. Answers are on a scale of 0-1, where 0 means "totally do not agree" and 1 stands for "totally agree". The last two columns of the table summarize the F-statistics and p values of the ANOVA test results.

For questions regarding rewards, trust, and the interpretation of a defective signal, I further conduct post-hoc pairwise comparisons between the treatments using Tukey HSD corrections for multiple tests. Subjects are more likely to reward their opponents in TP05 than in TN35 ($p = .005$), and in TP20 than in TN35 ($p = .035$). Subjcts trust their opponents more often in TP05 than in all treatments without communication ($p_{TP05vs.TN05} = .022$, $p_{TP05vs.TN20} < .001$, $p_{TP05vs.TN35} < .001$), more often in TP20 than in TN20 and TN35 ($p_{TP20vs.TN20} = .004$, $p_{TP20vs.TN35} < .001$), and more often in TP35 than in TN35 ($p_{TP35vs.TN35} = .009$). In terms of the interpretation of a defective, subjects in the pre-play communication treatments are more likely to attribute the defective signal to error if the opponent has promised to play C during the communication stage ($p_{TP05vs.TN05} = .001$, $p_{TP20vs.TN05} = 0.23$, $p_{TP35vs.TN05} = .013$, $p_{TP05vs.TN20} = .022$).

Abgrenzung

Das Projekt *Similarity and Consistency in Algorithm-Guided Exploration* ist ein Gemeinschaftsprojekt mit Fabian Dvorak, Ludwig Danwitz, Sebastian Fehrler, Lars Hornuf, Hsuan Yu Lin und Bettina von Helversen. Wir haben die Ideen für das Projekt zusammen durch viele Diskussionen entwickelt. Ich habe an den Simulationen vor unserer Datensammlung und später am Abschnitt zur Datenanalyse der Verhaltensergebnisse gearbeitet. Wir haben das Paper gemeinsam überarbeitet und bei *Management Science* eingereicht.

Das Projekt *Public Trust in Organizations* ist ein Gemeinschaftsprojekt mit Sebastian Fehrler und Volker Hahn. Die ursprüngliche Idee und das theoretische Modell stammen von Volker Hahn. Zusammen haben wir die Idee verfeinert und weiterentwickelt und einen ersten Entwurf erstellt. Nach der Vorregistrierung habe ich Daten gesammelt, unterstützt von Forschungsassistenten der Universitäten Konstanz und Bremen, und die Ergebnisse in unseren Entwurf integriert, um die erste Version des Papiers zu erstellen.

Das Projekt *Sustaining Cooperation With Correlated Information* ist ein gemeinsames Projekt mit Fabian Dvorak und Sebastian Fehrler. Wir haben die Idee zu diesem Projekt im ersten Jahr meines Doktoratsstudiums entwickelt. Ich habe das Experiment programmiert, alle Behandlungen im Labor durchgeführt und die aktuelle Version der Arbeit geschrieben.

Das Projekt *The Effect of Communication on Strategic Uncertainty in the Indefinitely Repeated Prisoners' Dilemma With Noise* ist mein unabhängiges Einzelautorenprojekt.