# Under Watchful Eyes: Observation by In-Group Leaders Promotes Intergroup Cooperation in Natural Groups

Juliane Koch* Lund University

Jonas Pilgaard Kaiser† Technical University of Berlin

October 22, 2025

**Abstract**

Promoting cooperation between different groups remains one of society's biggest challenges. We examine whether such intergroup cooperation can be fostered through observation by in-group leaders, using a lab-in-the-field experiment in Papua New Guinea. The effect of in-group observation is not clear ex ante, as individuals may both want to show themselves as cooperative ('cooperation motivation') and show that they prioritize their own group, for example, by only cooperating with members of their own group and not with others ('favoritism motivation'). We find that the cooperation motivation is stronger: Observation by the in-group leader increases the share of people who cooperate with members of other communities from 17 percent to 70 percent, thereby eliminating the in-group bias in cooperation. We relate this finding to a shared understanding among participants that intergroup cooperation is socially desirable. Our findings suggest that policymakers, communities, and organizations may leverage in-group observation to improve intergroup cooperation.

**JEL Classification:** C93, D91, H41, Q54

**Keywords:** Cooperation, in-group bias, leader observation, lab-in-the-field experiment, climate adaptation

---

* Corresponding author. Lund University, Department of Economics; juliane.koch@nek.lu.se. ORCID: 0009-0008-8697-7512. † Technical University of Berlin, Faculty of Economics and Management; kaiser@finance.tu-berlin.de. ORCID: 0000-0002-2554-8385.

The project was approved by the Institutional Review Board at Hamburg University (20.05.2023). The privacy rights of the participants have been observed and informed consent was obtained from all participants.

We especially thank Andreas Lange, Juan Camilo Cardenas, Gianluca Grimalda, Susann Adloff, Bertil Tungodden, Daniele Nosenzo, Roel van Veldhuizen, Dorothea Kübler, Claudia Schwirplies, Anna Hochleitner, and participants at the 2023 ESA World Meeting, the 2023 Workshop on Experimental Economics for the Environment, and the University of Gothenburg, the Swedish Agricultural University Uppsala, and the Beijer Institute Stockholm for helpful comments. We thank Freda Basse, Valentine Takahou, Boger Tsino, Serah Icavanamur, and Janette Kevi for excellent research assistance. We gratefully acknowledge the financial support of the German Research Foundation (DFG) within the Graduate Program Collective Decision-Making at the University of Hamburg, through the University of Hamburg, U4212GRK2001.

# 1 Introduction

Many of the most critical societal challenges require cooperation across groups. For example, nations must cooperate to mitigate climate change and make peace, opposing political parties must cooperate to ensure long-lasting improvements for society, and communities must work together to control diseases or respond to natural disasters. Yet, such intergroup cooperation is often hindered by the fact that it requires people to cooperate with someone who is more distant – geographically, socially, or culturally. An extensive literature on individuals' social identities shows that such cooperation is often difficult, as individuals may dislike or distrust members of other groups (cf. in-group bias, Tajfel et al., 1971; Turner et al., 1987; Chen and Li, 2009).[1] A key task for social scientists is therefore to find possible ways of promoting intergroup cooperation.

Within groups, cooperation often improves when actions are made observable to others (Samek and Sheremeta, 2014; Grimalda et al., 2016). One reason for this is that many individuals want to make a good impression on others (Bénabou and Tirole, 2006; Andreoni and Bernheim, 2009; Fiske, 2018) or receive social esteem from doing what is socially desirable (Krupka and Weber, 2013; Kimbrough and Vostroknutov, 2016). In between-groups interactions, however, little is known about the effect of making actions observable to other in-group members. This is critical, as individuals are often observed by members of their own groups, e.g., politicians within the same party, members of the same community, colleagues at the workplace, fellow students at the university, or family members. And it is not clear ex ante whether observation by in-group members hinders or improves intergroup cooperation, as there may be competing motivations at play: On the one hand, individuals may try to show to their in-group that they are cooperative, for example, by always cooperating with others ('cooperation motivation'). On the other hand, individuals may want to show that they prioritize their own group, for example, by only cooperating with members of their own group and not with members of other groups ('favoritism motivation').

This paper uses an experimental approach to study how observation by in-group leaders influences intergroup cooperation. Observation by leaders is particularly important, as leaders are not just ordinary group members: they embody the group's norms, have a higher social standing, and often serve as focal points for coordination (Baldassarri and Grossman, 2011; Acemoglu and Jackson, 2015; Kosfeld and Rustagi, 2015). To study the effect on intergroup cooperation, we conduct a lab-in-the-field experiment in Bougainville, Papua New Guinea. This field setting allows us to examine cooperation between naturally occurring groups of villages

[1] We follow Chen and Li (2009, p. 431) and define social identity as "a person's sense of self derived from perceived membership in social groups" (see also Tajfel, 1978).

from different geographical areas and ethnolinguistic groups. As we describe in Section 2, the setting is characterized by strong in-group biases between groups and a shared understanding of what is socially desirable within groups. The communities also have important informal hierarchies, which we exploit to test the effects of observation by in-group leaders.

In the experiment, participants make a decision in an incentivized, two-player Prisoner's Dilemma game that is modified to include exogenous risk. Participants face a binary choice of whether to cooperate or not, and the choice constitutes a social dilemma: The expected total earnings are greater if both participants cooperate, but any individual earns more by not cooperating. We vary in a $2 \times 2$ between-subjects design (i) whether participants interact with a member of their own community or with someone from another community, and (ii) whether participants are observed by their in-group leader or not.

Confirming our pre-registered hypotheses, we find that when participants act in private, they are 29 percentage points more likely to cooperate when they interact with a member of their own community than with a member of another community. This suggests that in-group bias can be a critical hindrance to cooperation across communities. However, observation by an in-group leader has a large, positive effect on intergroup cooperation, causing the share of individuals who cooperate to go from 17 percent to 70 percent. While observation also has a positive effect on cooperation with in-group members (from 46 percent to 59 percent), the effect is larger for intergroup cooperation, such that there is no in-group bias when people are observed. These results are corroborated by the participants' beliefs which indicate that intercommunity cooperation is socially desirable, suggesting that the 'cooperation motivation' is stronger than the 'favoritism motivation' in this setting. We discuss mechanisms in Section 6, where we i.a. use elicited beliefs to show that our results are more likely to be driven by treatment differences in preferences than beliefs about the behavior of the opposing participant.

This paper provides two important contributions to the literature on cooperation. Many studies seek to explain the determinants of cooperation, highlighting, for example, the role of trust, shared norms, and common objectives (Ostrom, 2009, 2010). Yet, individuals tend to be biased in favor of interacting with people from their own group, both due to intrinsic preferences and as a way to show group commitment (Turner, 1975; Chen and Li, 2009; Weisel and Böhm, 2015; Currarini and Mengel, 2016; Scheepers and Ellemers, 2019). As intergroup cooperation becomes increasingly important on both a local and a global scale, studies have examined ways of reducing this in-group bias. Some approaches focus on changing individuals' group categorizations, while others focus on changing the context of intergroup interaction.[2]

[2]For example, group categorizations can be altered by appealing to a common identity (Gaertner et al., 1989, 1993; Gaertner and Dovidio, 2000; Kershaw et al., 2021) or using cross-cutting cleavages in society (Coser, 1956; Dahl, 1956; Lipset, 1959; Lipset and Rokkan, 1967; Uğurlar et al., 2025). Other studies have added

The most closely related studies to ours are Charness et al. (2007), Charness and Rustichini (2011), and Habyarimana et al. (2007), who examine the effects of observation and punishment on intergroup cooperation.[3]

We are the first to study how in-group observation influences intergroup cooperation in a setting characterized by in-group bias. Importantly, we do so studying a natural field setting with groups that are intrinsically meaningful for the participants. As mentioned above, in-group observation is relevant because (i) it is common throughout society, and (ii) it is conceptually different from general observation, as it entails the tension between showing to one's group members a strong commitment to the group ('favoritism motivation') and showing oneself as a cooperative person ('cooperation motivation'). Our findings suggest that communities, organizations, and policymakers can harness in-group observation as a tool to foster intergroup cooperation – particularly in settings where cooperation between groups is socially desirable. This provides a practical pathway to reduce in-group bias and promote outcomes that benefit all groups.

A second contribution to the literature on cooperation is that we study the effect of observation in informal hierarchies in a field setting. Previous studies have examined how assigning leadership roles among participants in the lab can influence group behavior (Moxnes and van der Heijden, 2003; Levati et al., 2007; Oedzes et al., 2019; Yu and Kocher, 2023), for example, if a better-informed participant decides to lead by example (Potters et al., 2007). Some studies look instead at the importance of informal leaders in field settings. They show, for example, that central individuals are important for the spread of information (Banerjee et al., 2013), that the presence of a leader can reduce socially inappropriate punishment of defectors (Grimalda et al., 2016), and that members of congregations are influenced by learning how their religious leaders act towards members of other religions (Bauer et al., 2024). We extend this literature by investigating how being observed by an informal leader shapes interactions between groups. This is a central question, as informal hierarchies are widespread: They can be seen in parents or older siblings in families, a friend or neighbor who takes the lead in organizing events, or colleagues who coordinate teamwork. Our findings speak to the importance

communication (Charness et al., 2024), promoted previous contact between groups (Allport, 1954; Van Assche et al., 2023; Paolini et al., 2024; Braghieri et al., 2025), or used nudges (Dimant, 2024).

[3]Our study differs from Charness et al. (2007) and Charness and Rustichini (2011), for example, as they (i) study different groups of students that do not show in-group bias without observation, (ii) include a layer of home/guest framing to their setting, (iii) include face-to-face interaction with opposing participants, and (iv) include observability by the experimenter in the 'anonymous' treatment. Our study differs from Habyarimana et al. (2007), for example, as they (i) study general observation and not in-group observation, implying that the 'favoritism motivation' is muted, and (ii) combine observation with monetary punishment, thereby studying the effect of sanctions rather than image concerns from observation.

of leaders of informal hierarchies, as we show that being observed by the leader can promote prosocial behaviors and mitigate discrimination against other groups.

The remainder of the paper is structured as follows: Section 2 describes the field setting. Section 3 details the experimental design and the experimental procedure. We present a theoretical model and the pre-registered hypotheses in Section 4. The main results are presented in Section 5. Section 6 discusses further results and mechanisms, and Section 7 concludes. The online appendix includes the experimental instructions, more details on the study setting, a power analysis, as well as further results, tables, and figures.

## 2 Field Setting

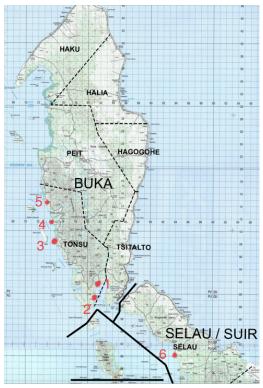
To study how observation by in-group leaders influences intergroup cooperation, we look for a setting with the following features: First, intergroup cooperation should be important for the groups, and cooperation should depend on individual members and not determined solely by group leaders. This implies that it is crucial that group members behave cooperatively. Second, observation by other group members should be feasible, and individuals should have a shared understanding of what observers want them to do. This ensures that observation affects individuals in a predictable and similar way. Third, there is a strong in-group bias, which means that it is challenging to achieve intergroup cooperation. As we explain in this section (with details in Appendix B), the Pacific island Bougainville in Papua New Guinea is characterized by these features and thereby provides a suitable field setting to study the interplay between observability and intergroup cooperation using natural groups (as opposed to, e.g., minimal groups in the lab).

Papua New Guinea (henceforth 'PNG') is the world's third-largest island country located in the south-western Pacific Ocean, and Bougainville is one of the country's autonomous islands, forming the largest island of the Solomon Islands' archipelago. PNG is renowned for its cultural and ethnic diversity with over 800 distinct languages and a vast array of customs and traditions (Reilly, 2008). According to World Bank indicators from 2025, 85 percent of the population lives in rural areas (The World Bank, 2025). Given the sparse infrastructure in rural areas, most villages are organized as independent communities with self-sufficient lifestyles.

Yet, cooperation across villages is increasingly important to Bougainville, which – similar to other Pacific islands – is heavily affected by climate change and the resulting natural hazards (Lang et al., 2020; Vousdoukas et al., 2023). Extreme temperatures and frequent climate disasters such as cyclones, droughts, and flooding endanger vital harvests (Wadey et al., 2017; Bourke, 2018). To protect against these climate hazards, villages rely on natural resources such

Figure 1: Field Setting


(a) Bougainville, Papua New Guinea




(b) Research Locations

Source: Google Maps, 2023

as mangrove trees and coral reefs that provide natural protection, for example, by reducing the wave intensity in case of coastal floodings. However, natural resources give rise to a critical social dilemma that goes across villages: The communities are better off with the natural resources intact, but individuals may exploit resources such as mangrove trees to generate higher income for themselves (Alongi, 2002; Wells et al., 2006; Sillanpää et al., 2024). To protect natural resources, it is crucial that the villages work together.

One issue with intercommunity cooperation in Bougainville, however, is the strong group categorization between villages. The combination of different ethnolinguistic groups, distinct cultural rituals, and the geographical disconnection between rural villages makes group categorizations occur naturally: People ascribe high values to their origin and village identity, while generally distrusting strangers (Morauta, 1974; Schram, 2015). This was enhanced through the Bougainville Civil War from 1988 to 1998, which further divided the population of Bougainville (Adamo, 2018). This makes Bougainville a highly relevant setting for studying group identities and in-group bias. In the experiment outlined in Section 3, we define in-group interactions as those occurring between members of the same village, whereas out-group interactions occur between participants from different villages.

Social life in traditional villages in PNG is characterized by hierarchical social structures and strong social norms. Specifically, communities are influenced by the traditional 'wantok' system that fosters strong rules of customs, norms, and kinship; a system that in some rural areas acts as a replacement for formal institutions (de Renzio, 2000; Nanau, 2011). Villages are led by a village leader (a so-called "Big Man" or "Big Woman"), who possesses informal

authority and imposes discipline to keep the traditional village life in order. The village leaders represent important nodes of the social network within villages, and they are important connectors when engaging with other villages. Often, they are the most respected people within a village, and other villagers seek to get along with them (Cochrane, 1970). Despite this hierarchical leader structure, there is a strong egalitarian community spirit in the villages, and many decisions and conflicts are resolved in regular community or mediation meetings (Boege, 2012). This way, problems are typically solved in a bottom-up communal way with the entire community serving as the primal actor (Regan, 2010). Hence, (i) the social context is very important to the villagers, (ii) the villagers have shared beliefs about what behaviors are desirable, and (iii) the behavior of the individual villagers matters greatly for the community. This makes Bougainville a particularly interesting setting for studying how observation by in-group leaders influences behavior.

Appendix B provides further details about the study setting, including a description of the sample used for the current experiment.

## 3 Experimental Design

In the following section, we describe the experiment that we use to study how observation by an in-group leader influences intergroup cooperation.[4] The experiment consists of four parts that participants complete in one session (see Figure 2 for an overview). First, participants make decisions in a modified Prisoner's Dilemma game with risk. Second, participants answer a battery of belief elicitation questions, including their beliefs about the actions of other participants and their beliefs about what the observer prefers that they do. Third, the risk in the Prisoner's Dilemma is resolved. Finally, participants complete a survey to provide additional control variables for the analysis. Experimental instructions are included in Appendix G. The experimental design, hypotheses, and analysis plan were preregistered in the AEA RCT registry (`https://www.socialscienceregistry.org/trials/11320`), and we document changes from the preregistration in Appendix A.

[4]Such causal insights would not be obtainable by simply observing the cooperative behavior of individuals, as such behavior is likely to be endogenous due to selection into (i) within- or between-group cooperation, (ii) situations with more or less observation by other in-group members, and (iii) groups that value between-group cooperation or in-group favoritism.

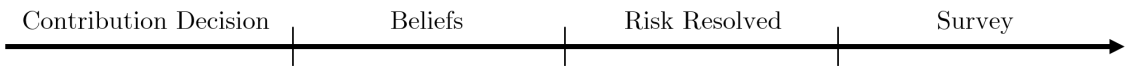



Figure 2: Timeline of the experiment

## 3.1 Modified Prisoner's Dilemma Game

For the first part of the experiment, participants play a version of the Prisoner's Dilemma game with risk. The Prisoner's Dilemma is a symmetric, simultaneous two-player game, in which each player decides whether to cooperate or defect. The game comprises a social dilemma, as the expected total earnings are higher when the players cooperate, but both players increase their own expected earnings by defecting. For this reason, the Prisoner's Dilemma provides a useful framework for studying when individuals want to cooperate with others. We include a risk component in the Prisoner's Dilemma to corroborate the framing we use in the experiment to ensure participants' understanding of the task (as described below).

### 3.1.1 Game

In the modified Prisoner's Dilemma game, each participant is endowed with 15 Kina (approximately 3.8 EUR) and randomly matched with one other anonymous participant. Participants learn that Nature makes a random draw between two states – a good state and a bad state, each occurring with a probability of 50 percent – and that they may contribute to a common fund to partly insure themselves and the other player against the bad state. In the bad state, both participants can end up losing their entire endowment; in the good state, both participants earn what they did not contribute. Before knowing the state, participants decide whether to cooperate by contributing 7 Kina of their endowment to a common fund or not. As contributions protect both players, a participant's payoff not only depends on their own choice but also on the choice of the other participant. Depending on the behavior of both players, there are three different levels of protection against the bad state: If none contributes, there is no protection, and both players lose their whole endowment. If only one contributes, both players are partially protected and lose 8 Kina. In this case, the player who contributed earns nothing ($15 - 7 - 8 = 0$), and the player who refrains from contributing earns 7 Kina ($15 - 8 = 7$). If both contribute, there is full protection, and both earn 8 Kina ($15 - 7 = 8$). If the good state is drawn, nothing is destroyed, and participants earn what they did not contribute.[5] Table 1 provides an overview of the possible payoffs.

[5]That is, if the bad state does not occur, the contributions that were made for protection are not returned to the participants. This reflects investments in any kind of insurance measure that deteriorates over time without having been used. Examples of such measures are manifold and include, for example, the building of unused or

Table 1: Payoff Matrices

| **Good State** | | |
|---|---|---|
| | Contribution | No Contribution |
| Contribution | 8;8 | 8;15 |
| No Contribution | 15;8 | 15;15 |
| **Bad State** | | |
| | Contribution | No Contribution |
| Contribution | 8;8 | 0;7 |
| No Contribution | 7;0 | 0;0 |
| **Expected Payoffs** | | |
| | Contribution | No Contribution |
| Contribution | 8;8 | 4;11 |
| No Contribution | 11;4 | 7.5;7.5 |

*Notes:* A random draw determines whether the good or bad state occurs (probability .5). In terms of expected payoffs, No Contribution is the dominating strategy although it leaves both players worse off than if both contributed.

As evident from the payoff matrices in Table 1, 'No Contribution' is a dominating strategy in the good state, and 'Contribution' is a weakly dominating strategy in the bad state. Yet, the participants do not know the state before making their decision. The strategy that maximizes expected payoffs is to not contribute to the common fund. That is, 'No Contribution' is a dominating strategy for a risk-neutral individual.

At the end of the experimental instructions, participants answer eight control questions that ensure that the participants understand the rules of the game. The participants had three tries to answer each control question correctly. If they did not, then the experimenter explained the instructions again until the participants answered the control question correctly. We show in Appendix D that the results are robust to excluding individuals who made multiple mistakes when answering the control questions.

After participants make their contribution decision and answer the battery of questions about beliefs (see Section 3.2), the random draw is made to determine whether the good or bad state occurs. Participants draw a card that indicates the state, and they do themselves to avoid any feelings of distrust towards the experimenter. This means that two cards are drawn

misplaced sea walls or other investments in disaster preparedness, buying vaccines for potential epidemics that do not break out, and disaster or military exercises for crises that do not occur.

for each pair of participants, and each participant draws the card that is payoff-relevant for themselves. The two-player pairs were matched ex post, such that participants who had drawn the bad state were randomly matched to another participant who had also drawn the bad state (and vice versa). While explaining this procedure to the participants, the experimenter has participants draw the card twice for practice, thereby internalizing the idea of a random draw.

### 3.1.2 Treatments

We apply a $2 \times 2$ between-subjects factorial design to obtain causal estimates of (i) how in-group bias influences whether participants cooperate, (ii) how observation by an in-group leader influences whether participants cooperate or not, and (iii) whether observation by an in-group leader has a positive or negative effect on interactions with out-group members, thereby potentially mitigating or enhancing in-group bias. Table 2 provides an overview of the treatments.

The first treatment dimension varies the origin of the other player. Specifically, participants are either matched with someone from their own village (in-group) or someone from another village (out-group). The out-group is chosen in such a way that while the villages are located in the same district (north-western part of the island), they belong to different constituencies with different local languages. That is, members of two different groups belong to different ethnolinguistic communities. Comparing behavior between the in-group and out-group conditions thus sheds light on the causal impact of in-group bias on cooperation.

The second treatment dimension varies whether participants are observed by their village leader when deciding whether to cooperate or not. The village leader is physically present in the room while the participant makes their decision (see Figure 3), and the village leader can see what the participant chooses. This introduces pressure to act in accordance with what the participant believes the village leader wants.[6] The village leaders do not make any active decision, and their earnings do not depend on the decisions made by the participants. Comparing behavior between the observed and private conditions sheds light on how observation by the in-group leader influences cooperation with members of one's in-group and out-group, respectively.

Previous studies with villagers in PNG demonstrate that although there is a shared understanding that free-riding is not desirable, there is also a shared belief that one should not gossip about or punish people for doing what is socially undesirable (Wiessner, 2020). We

[6]The advantage of having village leaders rather than other villagers act as observers is that the village leader is likely to be important for all participants (i.e., leaders are in everyone's 'reference network', Bicchieri, 2017). Thereby, we avoid the issue of some lab studies where the effect of observability may be dampened by participants perhaps not knowing or even caring about their observer.

confirm this for our sample using vignettes, as detailed in Appendix E.1. This indicates that observation by the village leader does not capture the effect of anticipated punishment or getting a worse reputation among the other villagers. Instead, the effect of being observed by the village leader likely stems from an increase in the social esteem (stigma) that participants receive from the village leader for doing what is socially desirable (undesirable), as modeled in Section 4.

Figure 3: Game Setting in the Field

(a) Observation (b) Private

Notes: The pictures show the conditions under which participants decide in the Observed and Private treatments, respectively. In both treatments, the experimenter first explains the game and then leaves the room when the decision is to be taken. In the observed treatment (a), the village leader is present when the participant makes their decision. In the private treatment (b), participants are alone when making their decision.

Table 2: Overview of Treatments

| | **Private** | **Observation** |
|---|---|---|
| **In-Group** | InPrivate | InObserve |
| **Out-Group** | OutPrivate | OutObserve |

### 3.1.3 Framing

As explained in Section 2, the villages of Bougainville have many important features for studying how in-group observation influences intergroup cooperation, using natural groups rather than artificial groups induced in the lab. Yet, this sample choice requires adaptations to the experimental design compared to the standard design used, for example, in lab studies with students. In particular, the large majority of villagers in Bougainville is not highly educated and does not regularly participate in economic experiments, for which reason it may be difficult for many villagers to understand abstract experimental games (Harrison and List, 2004; Alekseev et al., 2017). To enhance the participants' understanding, we instead provide context for the Prisoner's Dilemma, as suggested by e.g. Alekseev et al. (2017), by using the framing of climate change, which gives rise to a natural social dilemma that villagers are familiar with.[7] Specifically, participants decide whether to cooperate by "contributing to a climate fund", which serves to protect against a potential "natural disaster" that may otherwise destroy the earnings of the participants. Contributing to such funds is common practice in villages in Bougainville, as villagers regularly contribute money or time to community funds, which are used, for example, if the villagers need to repair parts of the community hall, church, or other public spaces. To explain the variation in group cooperation in the experiment, we vary whether the fund is a "community fund" (for in-group interaction) or a "district fund" (for out-group interaction).

By using this frame, we also mitigate the problem that participants may themselves create contexts for abstract experiments, which would challenge the controlled environment of the experiment. For example, Dufwenberg et al. (2011a) find that participants construct their own contexts to better deal with abstract terminology. Similarly, Harrison and List (2004, p. 1022) emphasize that for "abstract, context-free experiments [...] there is no control for the context that participants might themselves impose on the abstract experimental task."

Importantly, whereas the game is incentivized, the framing is hypothetical, and participants are aware of this. That is, the consequences of contributing to the climate fund reflect how such funds function in the field, namely to protect against potential climate damages. But all interactions are limited to the two players who are matched in the game, and the "natural disaster" is resolved by a random draw within the experiment.

In Appendix B, we provide additional information about the effects of climate change for Bougainville and the measures that communities take to alleviate the consequences thereof.

[7] Although context-neutral experimental instructions have long been favored by experimental economists, an increasing number of studies incorporate context in their instructions (for cooperation contexts, see e.g. Samuelson and Allison, 1994; Dufwenberg et al., 2011b; Eriksson and Strimling, 2014).

In Appendix E.2, we show that there is no interaction of the treatment effects with the participants' knowledge about environmental hazards, experiences with environmental hazards, attitudes towards climate change, perceived safety from climate disasters, or fear for food security due to environmental hazards. This indicates that while the framing of the current experiment was helpful in facilitating the understanding of the experiment for participants, the framing did not influence our reported treatment effects, thereby supporting the generalizability of the study's findings.

### 3.2 Belief Elicitation

After making their decision in the modified Prisoner's Dilemma game, participants in all treatments are asked a set of belief questions. The belief elicitation is unincentivized, as (i) using unincentivized questions is easier to understand for the participants, and it is therefore less likely to cause any confusion within this sample, and (ii) it reduces the time spent on the part of the experiment that is not related to making decisions in the Prisoner's Dilemma game. Participants report their beliefs about the behavior of the opposing player, their village leader, other participants from their village, and other participants from the other village. Subsequently, participants are asked about what they think the opposing player expects them to do, what the village leader wants them to do, and whether they fear that the village leader would dislike any contribution behavior that deviates from his expectation.[8] In addition, we ask participants about what they think that the village wants them to do. The full list of questions with the exact wording can be found in Appendix G.

### 3.3 Survey

After a random draw determines whether the good or the bad state was drawn in the Prisoner's Dilemma game, participants answer survey questions in three parts, concerning (i) demographics, (ii) social preferences, and (iii) topics related to climate change. First, the demographic questions ask about age, gender, education, type of work, and income. [9] Afterwards, participants answer questions related to their origin, as this may predict group identification with their village. The questions include mother tongue (open-ended text allowing for multiple

[8] As explained in Section 3.4, the experiment was carried out in Tok Pisin, which is the language that is taught in schools in the whole country. This language does not possess the same nuances as, for example, English for distinguishing between empirical and normative expectations (in the sense of Bicchieri, 2017).

[9] We adapt questions related to income to the setting of villages in Bougainville: As the majority of people does not earn money but live in a self-sufficient manner, often using barter, questions about monetary income provide insufficient information. To get a more complete picture of income, we therefore also ask about the money-recharging values on their cell phones and the frequency with which they use the internet.

answers), whether the participant was born in their current village, and how many years they have lived there in case they moved between villages. Next, the survey asks about community engagement both in terms of the type of engagement (time, money, or advice with multiple options possible) and the frequency of this engagement.

The second section of the survey relates to (social) preferences, taken from the World Value Survey (WVS), the General Social Survey (GSS), and the German Socio-Economic Panel (SOEP). First, participants answer trust questions to elicit both general trust and trust in strangers (2- and 4-point scales, respectively). Second, participants answer questions related to their perception of others' altruism and fairness.[10] Third, participants answer a question about their risk preferences (10-point scale, Dohmen et al., 2011). Then, participants state whether they think their reputation would improve the most by cooperating or not cooperating in the Prisoner's Dilemma.

The last part of the survey asks a battery of questions related to climate change. First, participants report what they think about when they hear the term 'climate change' (open-ended), and participants indicate whether they believe in climate change (Yes/No). Second, participants indicate whether they think a community climate fund would be beneficial for their village (Yes/No). Then, participants are presented with eight different natural hazards/environmental changes, and they answer for each phenomenon (i) whether they have heard about it (Yes/No), (ii) whether they have experienced it (Yes/No), and how important they think the phenomenon is for Bougainville (3-point Likert scale).[11] Finally, participants answer three vignettes related to climate adaptation measures. The vignettes provide further information about biases and preferences related to cooperation and climate change. The first vignette concerns in-group bias, as a person must decide in what villages to plant 20 trees to protect villages against flooding. The second vignette concerns luck egalitarianism in relation to climate hazards, as a woman earns more from a harvest than another woman only due to differences in luck (exposure to a storm). The third vignette concerns contributions to a climate community fund and how (un)fair it is for a person to free-ride on such contributions while benefitting from the fund. We examine the vignettes in detail in Appendix E.1.

[10] These questions were as follows: 'Do you think people are mostly looking out for themselves as opposed to trying to help each other' (perceived altruism, WVS); 'Do you think people would try to take advantage of you if they got a chance as opposed to trying to be fair' (perceived fairness variable, WVS).

[11] The phenomena were flooding/intense rainfall, tsunamis, earthquakes, sea level rise, intense drought, mountain erosion, change in weather patterns, and food scarcity/failed harvest.

### 3.4 Village Selection and Experimental Procedure

The lab-in-the-field experiment was conducted between April and May, 2023, in six small coastal communities in north-western Bougainville, Papua New Guinea. Data collection was completed in each village within two days to mitigate any discussion among villagers about the experiment. The village selection was concentrated to the north-western part of Bougainville (see Figure 1b). We chose this to ensure that all villages faced identical social dilemmas regarding the protection of the same natural resources, which ensured that the climate framing is comparable across all villages (see details in Appendix B). Furthermore, the villages are drawn from different districts within the same geopolitical subregion in Bougainville to ensure that the villages are comparable with respect to cultural variables while allowing for in-group and out-group distinctions.

Prior to data collection, pre-visits to each village were arranged to inform about and obtain consent for participation in the experiment.[12] Due to the hierarchical structure in the villages, we most often approached the village leader, who then asked the whole community whether they wished to take part in the research study. Once the participation was confirmed by the village representatives, the villagers were required to provide an up-to-date census to enable a random draw of participants. To recruit participants, households were drawn at random from the village census in each village, and one adult villager was randomly selected from each household. The resulting sample includes 60-70 adults from each of the six villages, yielding a total sample of 402 villagers.[13] In the sample, 52 percent were female, the mean age was 36 years, the average participant had received 8 years of schooling, participants on average earned a weekly income of 72 Kina (approximately 18.5 EUR), 79 percent reported to never use the internet, and 77 percent reported to attend church once a week. The full set of sample characteristics is provided in Appendix B.2.

At the start of each session, the research team (consisting of one international researcher and multiple local research assistants) was presented to the villagers. Then, the consent form for participation was read out loud and signed by all participants. The research was conducted by local research assistants in a pen-and-paper format in Tok Pisin, which is the main language taught in schools. Prior to data collection in the field, the research assistants

[12] In the pre-visits, we also explained that all participants would be paid after the data collection was completed in all villages. This procedure did not cause any concerns among participants.

[13] With this sample size, we expected based on power simulations to have a power of .8 to detect a decrease in cooperation (following from out-group interaction in private, cf. H1) of approximately 16 percentage points and an increase in cooperation (following from observation, cf. H2) of approximately 19 percentage points. We expected to have approximately .8 power to detect negative interaction effects of 15-27 percentage points and positive interaction effects of 12-26 percentage points (cf. H3). For details, see Appendix C.

were extensively trained to ensure a professional conduct and homogeneity in elicitation styles. All research assistants were trained to be able to conduct all parts of the experiment. The research assistants were not from any of the six sample villages and had not interacted with the study participants before.

The average duration of the experiment was approximately 50 minutes. Sessions took place in secluded spaces in community facilities of the village to ensure privacy during the decision process. All participants were paid 2 Kina as a show-up payment at the start of the experiment. Additionally, participants received on average a payout of 8.67 Kina from the modified Prisoner's Dilemma game, which was paid out in cash at the end of the data collection.

# 4 Theory and Hypotheses

In the following, we outline a simple model that provides a theoretical underpinning for our hypotheses about how in-group bias and observability influence cooperation. In doing so, we extend the model of social preferences by Charness and Rabin (2002) to incorporate both social identity (Chen and Li, 2009) and social esteem (Carpenter and Robbett, 2024).

## 4.1 Setup

**Utility.** We consider a population of individuals, who are matched into pairs and play a Prisoner's Dilemma game as described in Section 3. We assume that the utility of individual $i$ consists of (i) a weighted average of their own earnings and the earnings of the other player (Charness and Rabin, 2002), and (ii) utility derived from the social esteem or stigma that individual $i$ receives (Carpenter and Robbett, 2024). Formally, let $x_i$ ($x_j$) be the action of individual $i$ ($j$) of whether to cooperate, $C$, or defect, $D$. Individuals $i$ and $j$ act simultaneously in a one-shot interaction, and their resulting earnings are $\pi_i^s(x_i, x_j)$ and $\pi_j^s(x_i, x_j)$, respectively, where $s$ is the state that Nature draws, which is good with probability $p$ ($s = g$) or bad with probability $1 - p$ ($s = b$).

Furthermore, the utility of individual $i$ is influenced by the social esteem (stigma) they receive from cooperating or defecting (Krupka and Weber, 2013; Kimbrough and Vostroknutov, 2016; Barr et al., 2018), $N(x_i, \gamma_{ij})$.[14] Here, $\gamma_{ij}$ indicates whether the two individuals belong to the same group ($\gamma_{ij} = 1$) or to different groups ($\gamma_{ij} = 0$). Letting $N(\cdot)$ depend on $\gamma_{ij}$ implies that the social esteem (stigma) of cooperation and defection may depend on the social

[14] Alternatively, one could apply a signaling framework similar to Bénabou and Tirole (2006) to explain how an action $x_i$ may entail reputational benefits or costs for individual $i$.

identity of individual $j$. Specifically, we assume that it is not seen as better to cooperate with members of the out-group than with members of the in-group ($N(C,1) \geq N(C,0)$), and we assume that it is not seen as better to defect when interacting with members of the in-group than members of the out-group ($N(D,1) \leq N(D,0)$). In sum, the utility of individual $i$ in state $s$ is as follows:

$$u_i^s(x_i, x_j) = w_{ij}\pi_j^s(x_i, x_j) + (1 - w_{ij})\pi_i^s(x_i, x_j) + \nu_i N(x_i, \gamma_{ij}), \quad s = \{b, g\}, \tag{1}$$

where $w_{ij}$ is the weight that individual $i$ assigns to the payoff of individual $j$, and $\nu_i \geq 0$ is individual $i$'s sensitivity towards social esteem. The weight $w_{ij}$ depends on altruism and reciprocity, both of which may depend on the group identity of individual $j$:

$$w_{ij} = \lambda_i(1 + l_i\gamma_{ij}) + \rho_i(1 + r_i\gamma_{ij})\mathbb{1}_{x_j=D}, \tag{2}$$

where $\lambda_i \geq 0$ denotes the altruism of individual $i$ and $\rho_i \leq 0$ denotes the concerns for negative reciprocity ($\mathbb{1}_{x_j=D}$ equals 1 if $x_j = D$ and 0 otherwise). We allow altruism and reciprocity to depend on whether individuals $i$ and $j$ belong to the same group, following Chen and Li (2009). Specifically, we assume that $l_i > 0$, implying that individual $i$ values the earnings of members of the in-group more than the earnings of members of the out-group (Tajfel et al., 1971; Chen and Chen, 2011). Moreover, we assume $r_i \in (-1, 0)$, as Chen and Li (2009) show that individuals are more forgiving of misbehavior when interacting with members of the in-group compared to members of the out-group.

Individual $i$ chooses whether to cooperate or defect to maximize utility prior to knowing what state Nature draws. That is, the individual maximizes expected utility given as follows:

$$U_i(x_i, x_j) = p \cdot u_i^g(x_i, x_j) + (1 - p) \cdot u_i^b(x_i, x_j) \tag{3}$$

**Expected Earnings.** We assume that the expected earnings, $E_s[\pi_i^s(x_i, x_j)]$, give rise to a social dilemma in the absence of social preferences. For simplicity (and in accordance with the experiment outlined in Section 3), we let the expected earnings be symmetric and as follows: $E_s[\pi_i^s(C,C)] = R$, $E_s[\pi_i^s(C,D)] = S$, $E_s[\pi_i^s(D,C)] = T$, and $E_s[\pi_i^s(D,D)] = P$ with $T > R > P > S$ and $2R > T + S$. These assumptions imply that defection yields greater expected earnings regardless of the action of the other player, but the greatest total earnings are achieved when both cooperate.

**Incomplete Information.** We assume that individuals play a game of incomplete information, in which they know their own preference parameters but not those of the other player. Rather, individuals have a common prior: They know the probability distributions of all preference parameters. In particular, we assume that individual $i$'s type, $\theta_i \in \mathbb{R}^5$, is a vector

comprising the preference parameters in Equations 1 and 2. That is, individual $i$'s type is characterized as $\theta_i = (\lambda_i, l_i, \rho_i, r_i, \nu_i)$, and there is common knowledge about the probability distribution of types in the population, $f(\theta) : \mathbb{R}^5 \rightarrow [0, 1]$, which we assume to be continuous and integrable. We assume for simplicity that the probability distribution over preferences is independent across individuals in the population. Because we examine a static game of incomplete information, the relevant equilibrium concept in the following analysis is the Bayesian Nash Equilibrium.

### 4.2 Analysis

To examine equilibrium play, we first solve for when individual $i$ prefers to cooperate rather than defect. Individual $i$ does so when the expected utility of cooperating is greater than the expected utility of defecting. As individual $i$ knows their own type, the expectation is taken over what action individual $j$ takes. We denote by $q$ the probability that individual $j$ cooperates, implying that defection occurs with probability $1 - q$. Then, the optimal strategy for individual $i$ is a type-dependent strategy, where individual $i$ cooperates if and only if the following condition is satisfied:[15]

$$G(\lambda_i, \rho_i, l_i, r_i, \nu_i, \gamma_{ij}) \geq 1 - q, \tag{4}$$

Intuitively, the condition in Equation 4 gives rise to the predictions that, ceteris paribus, individuals who are more altruistic are more likely to cooperate ($G_{\lambda_i} > 0$), individuals with stronger concerns for negative reciprocity are less likely to cooperate ($G_{\rho_i} < 0$), and individuals who care more about social esteem are more likely to cooperate if they obtain more social esteem from cooperating than from defecting ($G_{\nu_i} > 0$ iff $N(C, \gamma_{ij}) > N(D, \gamma_{ij})$). Moreover, individual $i$ is more likely to cooperate with an in-group member ($G_{\gamma_{ij}} > 0$). If individuals $i$ and $j$ belong to the same group, then individual $i$ is more likely to cooperate when in-group altruism is higher ($G_{l_i} > 0$) and less likely to cooperate when in-group forgiveness is lower ($G_{r_i} < 0$).[16]

To form predictions for the treatment effects, however, we need to examine equilibrium

[15] For simplicity, we use a shorthand in Equations 4 and 5 by defining the following:

$$G(\lambda_i, \rho_i, l_i, r_i, \nu_i, \gamma_{ij}) \equiv \frac{T - R + (T - S)\lambda_i(1 + l_i\gamma_{ij}) + \nu_i[N(C, \gamma_{ij}) - N(D, \gamma_{ij})]}{T - R + P - S + (S - T)\rho_i(1 + r_i\gamma_{ij})}$$

[16] The condition in Equation 4 gives rise to a type-dependent, pure-strategy Bayesian Nash Equilibrium, where the outcome may deviate from the equilibrium of purely materially self-interested individuals, $(D, D)$, if the condition in Equation 4 is satisfied for some $\theta_i \in \Theta$. When Equation 4 holds with equality, individual $i$ is indifferent between cooperating and defecting and may thus mix strategies. Yet, this is inconsequential for the further analyses, as the continuity of $f(\theta)$ implies that indifference in Equation 4 occurs with probability zero.

play, as the condition for when individual $i$ cooperates also depends on the probability with which individual $j$ cooperates. Due to symmetry, the strategy of individual $j$ is also a type-dependent threshold strategy. This means that we can redefine $q$ as the probability with which individual $j$ belongs to a type for whom Equation 4 is satisfied. Specifically, denote by $A \subset \Theta$ the subset of types for whom Equation 4 is satisfied. Then, $q = Pr(x_j = C) = Pr(\theta_j \in A)$, and we can rewrite the condition for when individual $i$ cooperates:

$$G(\lambda_i, \rho_i, l_i, r_i, \nu_i, \gamma_{ij}) \geq 1 - \int_A f(\theta) d\theta \tag{5}$$

For the treatment effects, we examine first the influence of in-group bias on cooperation when individuals act in private either with a member of their own group ($\gamma_{ij} = 1$) or a member of another group ($\gamma_{ij} = 0$). As seen in Equation 5, there is more cooperation on average when individuals interact with members of their own group: When $\gamma_{ij} = 1$, concerns for altruism increase, concerns for negative reciprocity decrease, and the social esteem of cooperating instead of defecting weakly increases, all of which leads to $G_{\gamma_{ij}} > 0$. In addition, Equation 5 is satisfied for a larger subset of types $\theta$ when $\gamma_{ij} = 1$ compared to when $\gamma_{ij} = 0$, implying that $A|_{\gamma_{ij}=0} \subset A|_{\gamma_{ij}=1}$ and that the right-hand side of Equation 5 decreases. Together, this leads Equation 5 to be satisfied for more individuals, thereby providing the first hypothesis:

**Hypothesis 1** *More participants cooperate in InPrivate than in OutPrivate.*

Second, we examine how observability influences cooperation for individuals who interact with members of their in-group. We assume that a change in observability influences the sensitivity towards social esteem or stigma. We model this through a change in $\nu_i$ from the private ($\nu_i^p$) to the observed setting ($\nu_i^o$), with $\nu_i^p < \nu_i^o$. We expect that individuals in our setting receive more esteem from cooperating than defecting (Boege, 2013; Autonomous Bougainville Government, 2023), which implies that $G_{\nu_i} > 0$, and observability leads to an increase in the left-hand side of Equation 5. Similar to the case for H1, we have that if individuals receive more esteem from cooperating than defecting, then Equation 5 is satisfied for a larger subset of types for $\nu_i^o$ than for $\nu_i^p$, implying that $A|_{\nu_i=\nu_i^p} \subset A|_{\nu_i=\nu_i^o}$ and that the right-hand side of Equation 5 decreases with observability. We thus reach our second hypothesis:

**Hypothesis 2** *More participants contribute in InObserve than in InPrivate.*

Finally, we look at the interaction between observability and in-group bias to examine how observation by an in-group leader may influence intergroup cooperation and in-group bias. As explained above, we assume that the utility of individual $i$ consists of (i) a weighted average of

the earnings of individuals $i$ and $j$ and (ii) utility derived from the social esteem or stigma of the action. Because (i) does not depend on observability, any effect of observability on intergroup cooperation and in-group bias works through (ii). Importantly, the effect of observation on out-group interactions depends on whether individuals receive social esteem or stigma from cooperating with members of the out-group. For example, if intergroup cooperation is socially desirable ('cooperation motivation'), then observation should increase cooperation with out-group members; however, if it is more socially desirable to only cooperate with in-group members ('favoritism motivation'), then observation should decrease cooperation with out-group members. We do not know the relative strengths of the cooperation and favoritism motivations ex ante, but we elicit beliefs about social desirability in the experiment and return to these in Section 5.3.

Further, given the expected in-group bias (H1) and positive effect of observation on in-group cooperation (H2), we are interested in the extent to which observation may reduce or enhance in-group bias. A reduction occurs when observation increases cooperation more with out-group members than with in-group members. A necessary condition for this is the following: If cooperation is socially desirable among in-groups, i.e., $N(C, \gamma_{ij} = 1) > N(D, \gamma_{ij} = 1)$, then individuals must gain more social esteem from cooperating than defecting when interacting with out-groups, $N(C, \gamma_{ij} = 0) > N(D, \gamma_{ij} = 0)$. This condition is not sufficient, however, as the effect on in-group bias depends on the relative esteem in in-group and out-group interactions as well as potential ceiling effects: For example, if sufficiently many individuals cooperate with in-groups in private, then observability could have a limited effect on in-group cooperation compared to out-group cooperation, thereby reducing in-group bias. Thus, our third hypothesis involves competing predictions and is as follows:

**Hypothesis 3**

*a. Observability increases in-group bias in contribution levels:*
*(OutObserved-InObserved) > (OutPrivate-InPrivate)*

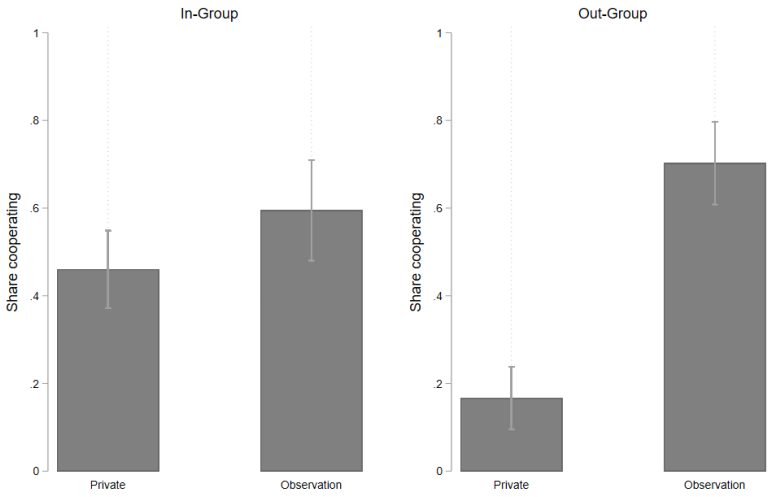
*b. Observability decreases in-group bias in contribution levels:*
*(OutObserved-InObserved) < (OutPrivate-InPrivate)*

# 5 Results

In this section, we present the results on how the identity of the other player, observation, and the combination of the two influence contributions in the experiment. Throughout, we use logit regressions for the primary tests, and we estimate the regressions with (i) no controls, (ii) demographic controls, and (iii) demographic and attitudinal controls. Specification (iii)

is our preferred specification, as we expect the control variables to capture variation in the contribution rates other than what is caused by the treatments. Note that some control variables were not answered by all participants, and this leads to a slightly lower number of observations in specifications (ii) and (iii). In this section, we report the analysis using the maximal number of participants for each specification; in Appendix D, we show that the results are robust to including only the participants who answer all control questions. We use the nonparametric chi-squared and Fisher's exact tests for robustness when appropriate, and we report other robustness checks in Section 5.4. Throughout, we report significance levels with two-sided tests. We discuss exploratory findings in Section 6. We list any deviations from the preregistration in Appendix A.

Figure 4: Contribution Behavior Across Treatments




Notes: This figure shows the share of participants contributing for each treatment. Error bars indicate 95 percent confidence intervals.

## 5.1 H1: In Private, Participants Cooperate Less With Members of the Out-Group

First, we report the results related to H1, which states that more participants cooperate in InPrivate than in OutPrivate. As illustrated in Figure 4, this is indeed the case with 46 percent of participants contributing in InPrivate as opposed to only 17 percent in OutPrivate. In logit regressions, this effect is statistically significant in all specifications (all $p's < .001$, cf. Table F.20), and the effect is robust (chi-squared: $p < .001$; Fisher's exact: $p < .001$). Accounting for all control variables (Column (3) in Table F.20), participants who decide in private are

on average 24 percentage points less likely to cooperate when interacting with a member of another village than when interacting with a member of the same village. In terms of practical significance, this effect is substantial, and it is similar to e.g. the level of in-group bias of 31 percentage points that Ahmed (2007) find using a Prisoner's Dilemma game. Supporting H1, we thus conclude the following:

**Result 1** *When participants decide in private, they are more likely to cooperate when interacting with a person from their own village than from another village.*

In Appendix E.1, we show that the treatment effect is larger among villagers with stronger in-group bias (as measured in vignettes), indicating that the treatment variation indeed captures the effect of in-group bias. We discuss in Section 6.1 that this in-group bias seems to stem from the participants' preferences for cooperating with in-group members and not from differences in beliefs about whether the other player cooperates.

### 5.2 H2: Observation Increases Cooperation in In-Groups

Second, we report the results related to H2, which states that more participants cooperate in InObserve than in InPrivate. From Figure 4, this hypothesis seems to be confirmed: 54 percent of participants in InObserve cooperate compared to 46 percent in InPrivate. In logit regressions, this effect is only marginally significant without controls ($p = .064$); however, it becomes statistically significant when adding demographic controls ($p = .039$) and highly significant when also including attitudinal controls ($p = .006$), cf. Table F.21. With nonparametric tests, the effect is also marginally significant (chi-squared: $p = .067$; Fisher's exact: $p = .079$), which is expected given the parametric results, as neither the chi-squared nor the Fisher's exact test accounts for control variables.

In terms of practical significance, the effect of observation is substantial: When accounting for all control variables (Column (3) in Table F.21), participants interacting with a person from their own village are on average 22 percentage points more likely to cooperate when they are being observed compared to when they decide in private. This effect size also resembles that of Grimalda et al. (2016), who find that observation increases cooperation in a Prisoner's Dilemma game by 17 percentage points. Supporting H2, we sum up this result as follows:

**Result 2** *When participants interact with a member of their own village, they are more likely to cooperate when they are observed by the village leader compared to when they decide in private.*

### 5.3 H3: Observation Mitigates the Effect of In-Group Bias

Third, we report the results related to H3, which examines whether observation by the village leader influences the in-group bias demonstrated in Section 5.1. Following the theoretical model presented in Section 4, a necessary condition for observation to reduce in-group bias is that individuals receive more social esteem from cooperating than defecting when interacting with an out-group member. We first examine whether this is satisfied by looking at beliefs the participants have about what the village leader and the village prefer that the participant does.

In all treatments, a majority of participants believe that the village leader wants them to contribute (68-79 percent across treatments, binomial tests: all $p's < .001$) and that their village wants them to contribute (between 64-78 percent, binomial tests: all $p's < .001$). None of these beliefs depend on either of the treatment dimensions (logit, chi-squared test, Fisher's exact test: all $p's > .186$). This indicates that there is a shared belief that contribution is socially desirable, also with members of the out-group. Thus, we anticipate that observation also increases contribution with out-group members, and we therefore turn now to examine whether observation influences in-group bias.

We find that observation increases contribution levels in out-group interactions from 17 percent to 70 percent, which is significant for all specifications (all $p's < .001$) and robust (chi-squared: $p < .001$; Fisher's exact: $p < .001$). Observation has an even greater effect on out-group interactions, leading to a reduction in in-group bias in contribution levels: The in-group bias without observation corresponds to a difference in support for contribution of 29 percentage points (InPrivate – OutPrivate). Opposingly, there is no in-group bias with observation, and the sign actually reverses to -11 percentage points (InObserve – OutObserve), which is statistically insignificant (all $p's > .148$, cf. Table F.23). We discuss this in Section 6.2. As seen in Table 3, the reduction in in-group bias is highly statistically significant in all specifications (all $p's < .001$). And with a decrease in the effect of out-group affiliation of 34 percentage points (Column (3) in Table 3), we conclude that the effect is not only of statistical but also practical significance. Supporting H3b, we thus sum up this result as follows:

**Result 3** *Observation by the village leader increases cooperation with out-group members substantially and eliminates the in-group bias in cooperation.*

### 5.4 Robustness

In Appendix D, we show that the results are robust to (i) using Linear Probability Models instead of logit regressions, (ii) excluding participants who needed many attempts to answer

Table 3: H3: Interaction Effect of In-Group Bias and Observation on Contributions

| | (1) | (2) | (3) |
|---|---|---|---|
| Out-group | -0.28*** | -0.28*** | -0.24*** |
| | (0.05) | (0.05) | (0.06) |
| Observation | 0.12* | 0.14** | 0.21*** |
| | (0.07) | (0.07) | (0.08) |
| Observation × Out-group | 0.40*** | 0.38*** | 0.34*** |
| | (0.07) | (0.08) | (0.10) |
| Baseline Cooperation Rate | .46 | .50 | .43 |
| Demographic Controls | No | Yes | Yes |
| Attitudinal Controls | No | No | Yes |
| Observations | 402 | 379 | 315 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The demographic controls are age, gender, years of schooling, weekly income, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. A table with controls can be found in Appendix F.22. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the control questions correctly, (iii) excluding in all specifications the participants who do not answer all demographic and attitudinal questions, (iv) excluding the 10 participants (2.5 percent of the sample) who were flagged by experimenters for potential problems with the experimental procedure, (v) using alternative variables as proxies for participants' income, and (vi) adding village fixed effects.

In Appendix D.8, we also show that we find no evidence of clustering across villages in any of the specifications used to test the hypotheses. Moreover, the results are robust to using both cluster-robust standard errors and score wild cluster bootstrap (Kline and Santos, 2012).

We next examine the robustness of our findings to correcting for multiple hypotheses. Specifically, we test three pre-registered hypotheses, which implies that the overall type I error rate – the probability of finding an effect when there is in fact no effect – is larger than the significance level of .05. To correct for this, we compute the adjusted $p$-values when controlling for the family-wise error rate (FWER), which is the probability of falsely rejecting at least one null hypothesis among all hypotheses in a family of tests (all three hypotheses in our case). To do so, we use the Holm-Šidák procedure (Šidák, 1967; Holm, 1979), and we use the $p$-values from our preferred specification that includes all control variables. We find that all treatment

effects remain highly statistically significant (H1: $p = .001$, H2: $p = .009$, H3: $p = .004$).

# 6 Beliefs and Mechanisms

The preceding analysis tested pre-registered hypotheses and showed that (i) in-group bias influences cooperation when individuals act anonymously, (ii) observation by an in-group leader increases cooperation levels with in-group members, and (iii) observation by an in-group leader increases cooperation levels even more with out-group members, effectively mitigating in-group bias.

In what follows, we take an exploratory approach to the data. First, we demonstrate that the treatment effects are likely to be driven by changes in preferences rather than beliefs about what the other player does. Second, we discuss why observation has a greater effect on cooperation with out-group than in-group interactions in our setting.

In Appendix F.2, we discuss the results of the vignettes that document that (i) in-group bias in vignettes predict responsiveness to the group condition, (ii) there is a shared belief that free-riding on contributions to a common fund is wrong, and (iii) villagers disapprove of gossiping about or punishing other villagers. Then, we show that the framing of the experiment is unlikely to have influenced our reported treatment effects. We also discuss further the social expectations in the modified Prisoner's Dilemma game, showing that there are different expectations about the behavior of the in-group leaders compared to the other villagers. We then document that observation reduces the in-group bias more for men than for women. This could be explained by men exhibiting a greater in-group bias when acting in private, and we discuss this in relation to the different roles of men and women in the villages. Finally, we show that there are no heterogeneous treatment effects across age, income, education, or years of schooling.

## 6.1 The Role of Beliefs and Preferences

A large literature documents that many people are conditional cooperators, deciding to cooperate with other people only if they believe that the others cooperate (Fischbacher et al., 2001, 2012). Thus, the treatment effects from Section 5 may arise from two different sources: First, individuals may have different preferences across the different treatments, e.g., due to parochial altruism or a desire to obtain social esteem. Second, individuals may expect that the behavior of the other player differs, thereby influencing participants' behavior if they are conditional cooperators. We document in this section that the second explanation is less likely to be driving the results of this paper.

The current design provides no causal explanation for how expectations about the behavior of the other player influences the willingness to cooperate. Yet, correlational evidence shows a positive association between contributing and believing that the other player will contribute. Across all treatments, participants who answer that they think the other player contributes are 15 percentage points more likely to contribute themselves compared to participants who think the other player defects. (logit, $p < .004$, cf. Table F.24). Yet, this relation is not causal, and it is possible that the causality goes in the other direction, e.g., if a false-consensus effect makes participants believe that the other player behaves similarly to themselves (Ross et al., 1977).

To test whether the treatment influences behavior through beliefs, we examine whether there are treatment differences in beliefs about what the other player does. To do so, we conduct analyses analogous to those in Section 5 with the expected behavior of the other player as dependent variable. Looking first at the influence of the group affiliation of the other player, we find that participants become 10 percentage points less likely to expect that the other player contributes when the other player belongs to an out-group, but this effect is not statistically significant ($p = .184$, cf. Table F.25). Similarly, we find no effect of observation on the participants' expectations about what the other player does ($p = .229$, cf. Table F.26). If anything, observability *reduces* the probability that a participant expects the other player to contribute from 53 to 44 percent. Finally, we find in a logit regression with all treatments that neither the direct effects nor the interaction of treatments are statistically significant (cf. Table F.27). These results suggest that the treatments have little to no effect on the participants' expectations about what the other player does, suggesting that the treatments work through something other than expectations.

Another way to examine whether treatment effects go through expectations about the other player is to include beliefs about the actions of the other player as a control variable in the logit regressions estimated in Section 5. If the treatment effects were mediated through beliefs, then one would expect that adding beliefs as a control variable would reduce the partial effects of the treatments. Yet, as shown in Tables F.28-30, including beliefs in the treatments does not alter the treatment effects, neither in terms of their approximate size nor statistical significance.

In sum, the data provides little support for the treatment effects to run through average beliefs about what the other player does.[17] In contrast, we show in Appendix E.1 that varying

[17]The current experiment only elicits binary beliefs about the actions of others (Contributing vs. Not contributing), and it therefore does not speak to uncertainty in beliefs about the response of the other player. Uncertainty could, e.g., matter for participants' decisions if they seek to lower the probability with which they are in a situation where they contribute and the other person does not contribute (cf. 'sucker aversion',

the group affiliation of the other player has a significantly larger effect among villagers with stronger preferences for benefitting the in-group (as measured in vignettes). This suggests that treatment differences are more likely to arise due to different preferences depending on who the participants interact with and whether they are observed. This emphasis on preferences resembles the results found in e.g. De Cremer and Van Vugt (1999), but deviates from the results found in e.g. Habyarimana et al. (2007).

### 6.2 The Greater Effect of Observation Among Out-Groups

As we show in Section 5.3, observation in our setting has a stronger effect on cooperation when people interact with out-group members than with in-group members. One possible reason for this is mechanical: Because contribution levels are lower for out-group interactions without observation (46 percent in InPrivate vs. 17 percent in OutPrivate), there is a greater scope for observation to increase cooperation with out-groups. Yet, the mechanical explanation seems insufficient: Observation can still increase in-group cooperation (cf. Result 2), and it even causes slightly *more* people to cooperate with out-groups than in-groups (59 percent in InObserve vs. 70 percent in OutObserved, all $p's > .148$).

A second possible explanation is that the signaling value of cooperating may be greater for interactions with out-group members compared to in-group members. From a Bayesian perspective, the reputation that a participant has with the village leader will be a combination of (i) the village leader's prior beliefs about the participant and (ii) the signal that the participant sends with their action in the experiment. In interactions with in-group members, prior beliefs about cooperation are likely well-established: People frequently engage with their in-group, and their past behavior, reputation, and commitment to the group are known within their social networks. Thus, a single cooperative act within the experiment is just one of many signals contributing to a broader, stable impression of the participant's social behavior. In contrast, interactions with out-group members tend to be more sporadic, and prior beliefs about the participant's cooperative tendencies are less certain. This would lead the village leader to put more weight on new signals, making cooperation with out-group members more valuable for the participants.

This explanation is corroborated by previous literature that demonstrates how past information influences the effect of observability. For example, Adloff and Pondorfer (2024) study the role of social distance between participants and observers in a Dictator Game. They find

Wilkinson-Ryan, 2008; Bougherara et al., 2009). For instance, if participants who believe that the other player contributes become more uncertain about this when facing a member of another village, this could lead to lower contribution rates when facing an out-group member compared to an in-group member.

that observation by distant observers has greater effects on behavior and argue that while distant observers learn much from the behavioral signal in the experiment, "close observers are likely well informed about the decision-makers personality and thus, hardly learn anything new" (p. 4). Similarly, Exley (2018) find that individuals become less responsive to image concerns when observers know their past histories of volunteering. In what she calls "the Reputations Effect", she argues that "if individuals' reputations about past volunteer behavior are public, choosing to volunteer may be less informative about their prosocial tendencies" (p. 1).

# 7 Conclusion

This paper addresses the key challenge of how to promote intergroup cooperation. Using a lab-in-the-field experiment in Bougainville, Papua New Guinea, we first show that in-group bias is a hindrance to intergroup cooperation, as individuals are less likely to cooperate with out-group members than with in-group members. Yet, when individuals are observed by their in-group leader, they are much more likely to cooperate with out-group members. The effect is larger for intergroup cooperation than for cooperation with members of one's own group, and in-group observation therefore eliminates the in-group bias.

Our findings suggest that communities, organizations, and policymakers may leverage in-group observation to improve intergroup cooperation when group members have a shared understanding that intergroup cooperation is socially desirable. This insight is broadly applicable and can help foster cooperation between groups. It may be particularly important in settings with informal hierarchies, as these rely rather on social pressure than on formal rules, for which reason many standard policy tools may not be available. This is the case in many environments where people need to work together across groups, for example, to mitigate negative consequences of climate change, eradicate diseases, ensure public safety, and improve infrastructure. In such situations, cooperation is often more difficult to sustain with out-groups (see, e.g., Forsyth, 2013, for the case of community-based climate adaptation, where initiatives tend to stay local rather than to scale across regions). Our study shows that making actions observable to others within the same community can improve intercommunity cooperation.

Nonetheless, there is potential to broaden the scope and enhance the generalizability of our findings. First, our study focuses on a setting where intercommunity cooperation is socially desirable, but it does not explore the underlying cultural or historical factors that shape these views, and we consider this as an important direction for future research. Second, the study was conducted in one cultural and geographical context, and this may limit the external validity of the findings. Many researchers have called for studies that examine non-WEIRD

populations (Henrich et al., 2010), which is what we do in this paper, but it is important for future studies to also replicate the current findings in other populations. Lastly, while we show that in-group observation increases intergroup cooperation, we do not examine how this effect impacts future interactions between the groups. An interesting avenue for future research is to examine the conditions under which shifts in intergroup cooperation persist over time.

# References


Acemoglu, D. and M. O. Jackson (2015): "History, Expectations, and Leadership in the Evolution of Social Norms," *The Review of Economic Studies*, 82, 423–456.

Adamo, A. (2018): "A Cursed and Fragmented Island: History and Conflict Analysis in Bougainville, Papua New Guinea," *Small Wars & Insurgencies*.

Adloff, S. and A. Pondorfer (2024): "Social Image in Context: The Role of Social Norms and Social Networks," Tech. rep., Working Paper.

Ahmed, A. M. (2007): "Group Identity, Social Distance and Intergroup Bias," *Journal of Economic Psychology*, 28, 324–337.

Alekseev, A., G. Charness, and U. Gneezy (2017): "Experimental Methods: When and Why Contextual Instructions Are Important," *Journal of Economic Behavior & Organization*, 134, 48–59.

Allport, G. W. (1954): *The Nature of Prejudice*, Cambridge, MA: Perseus Books.

Alongi, D. M. (2002): "Present State and Future of the World's Mangrove Forests," *Environmental conservation*, 29, 331–349.

Andreoni, J. and B. D. Bernheim (2009): "Social Image and the 50–50 Norm: A Theoretical and Experimental Analysis of Audience Effects," *Econometrica : journal of the Econometric Society*, 77, 1607–1636.

Autonomous Bougainville Government (2023): *Bougainville Long Term Vision 2052*, Aoutonomous Bougainville Government.

Baldassarri, D. and G. Grossman (2011): "Centralized Sanctioning and Legitimate Authority Promote Cooperation in Humans," *Proceedings of the National Academy of Sciences*, 108, 11023–11027.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013): "The Diffusion of Microfinance," *Science*, 341, 1236498.

Barr, A., T. Lane, and D. Nosenzo (2018): "On the Social Inappropriateness of Discrimination," *Journal of Public Economics*, 164, 153–164.

Bauer, M., J. Chytilová, and E. Ochieng (2024): "Preference Transmission within Churches: Religious Leaders and Clusters of (In)Tolerance," Working paper, The Center for Economic Research and Graduate Education - Economics Institute, Prague.

Bénabou, R. and J. Tirole (2006): "Incentives and Prosocial Behavior," *American Economic Review*, 96, 1652–1678.

Bicchieri, C. (2017): *Norms in the Wild*, New York, NY: Oxford University Press.

Boege, V. (2012): "Hybrid Forms of Peace and Order on a South Sea Island: Experiences from Bougainville (Papua New Guinea)," in *Hybrid Forms of Peace: From Everyday Agency to Post-Liberalism*, Springer, 88–106.

——— (2013): "Bougainville Report. Project: Addressing Legitimacy Issues in Fragile Post-Conflict Situations to Advance Conflict Transformation and Peacebuilding," Tech. rep., Berghof Foundation.

Bougherara, D., S. Costa, G. Grolleau, and L. Ibanez (2009): "Dealing with Aversion to the Sucker's Payoff in Public Goods Games," *Economics Bulletin*, 29, 3194–3202.

Bourke, R. M. (2018): "Impact of Climate Change on Agriculture in Papua New Guinea," *Climate Change: Our Environment, Livelihoods, and Sustainabilit*, 35–50.

Braghieri, L., P. Schwardmann, and E. Tripodi (2025): "Talking across the Aisle," *Unpublished*.

Carpenter, J. and A. Robbett (2024): "Measuring Socially Appropriate Social Preferences," *Games and Economic Behavior*, 147, 517–532.

Charness, G., G. D. Bartolomeo, and S. Papa (2024): "A Stranger in a Strange Land: Promises and Identity," *Games and Economic Behavior*, 144, 13–28.

Charness, G. and M. Rabin (2002): "Understanding Social Preferences with Simple Tests," *The Quarterly Journal of Economics*, 117, 817–869.

Charness, G., L. Rigotti, and A. Rustichini (2007): "Individual Behavior and Group Membership," *American Economic Review*, 97, 1340–1352.

Charness, G. and A. Rustichini (2011): "Gender Differences in Cooperation with Group Membership," *Games and Economic Behavior*, 72, 77–85.

Chen, R. and Y. Chen (2011): "The Potential of Social Identity for Equilibrium Selection," *American Economic Review*, 101, 2562–89.

Chen, Y. and S. X. Li (2009): "Group Identity and Social Preferences," *American Economic Review*, 99, 431–457.

Cochrane, G. (1970): *Big Men and Cargo Cults*, Oxford: Clarandon Press.

Coser, L. A. (1956): *The Functions of Social Conflict*, Clencoe, IL: The Free Press.

Currarini, S. and F. Mengel (2016): "Identity, Homophily and in-Group Bias," *European Economic Review*, 90, 40–55.

Dahl, R. A. (1956): *A Preface to Democratic Theory*, Chicago: The University of Chicago Press.

De Cremer, D. and M. Van Vugt (1999): "Social Identification Effects in Social Dilemmas: A Transformation of Motives," *European Journal of Social Psychology*, 29, 871–893.

de Renzio, P. (2000): "Bigmen and Wantoks: Social Capital and Group Behaviour in Papua New Guinea," *QEH Working Papers*, qehwps27, Queen Elizabeth House, University of Oxford.

Dimant, E. (2024): "Hate Trumps Love: The Impact of Political Polarization on Social Preferences," *Management Science*, 70, 1–31.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011): "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences," *Journal of the European Economic Association*, 9, 522–550.

Dufwenberg, M., S. Gächter, and H. Hennig-Schmidt (2011a): "The Framing of Games and the Psychology of Play," *Games and Economic Behavior*, 73, 459–478.

Dufwenberg, M., P. Heidhues, G. Kirchsteiger, F. Riedel, and J. Sobel (2011b): "Other-Regarding Preferences in General Equilibrium," *The Review of Economic Studies*, 78, 613–639.

Eriksson, K. and P. Strimling (2014): "Spontaneous Associations and Label Framing Have Similar Effects in the Public Goods Game," *Judgment and Decision Making*, 9, 360–372.

Exley, C. (2018): "Incentives for Prosocial Behavior: The Role of Reputations," *Management Science*, 64, 2460–2471.

Fischbacher, U., S. Gächter, and E. Fehr (2001): "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment," *Economics Letters*, 71, 397–404.

Fischbacher, U., S. Gächter, and S. Quercia (2012): "The Behavioral Validity of the Strategy Method in Public Good Experiments," *Journal of Economic Psychology*, 33, 897–913.

Fiske, S. T. (2018): *Social Beings: A Core Motives Approach to Social Psychology*, Hoboken, NJ: John Wiley & Sons, Inc.

Forsyth, T. (2013): "Community-Based Adaptation: A Review of Past and Future Challenges," *WIREs Climate Change*, 4, 439–446.

Gaertner, S. and J. Dovidio (2000): *Reducing Intergroup Bias: The Common Ingroup Identity Model*, Philadelphia: Psychology Press.

Gaertner, S. L., J. F. Dovidio, P. A. Anastasio, B. A. Bachman, and M. C. Rust (1993): "The Common Ingroup Identity Model: Recategorization and the Reduction of Intergroup Bias," *European Review of Social Psychology*, 4, 1–26.

Gaertner, S. L., J. Mann, A. Murrell, and J. F. Dovidio (1989): "Reducing Intergroup Bias: The Benefits of Recategorization," *Journal of Personality and Social Psychology*, 57, 239–249.

Grimalda, G., A. Pondorfer, and D. P. Tracer (2016): "Social Image Concerns Promote Cooperation More than Altruistic Punishment," *Nature Communications*, 7, 12288.

Habyarimana, J., M. Humphreys, D. N. Posner, and J. M. Weinstein (2007): "Why Does Ethnic Diversity Undermine Public Goods Provision?" *The American Political Science Review*, 101, 709–725.

Harrison, G. W. and J. A. List (2004): "Field Experiments," *Journal of Economic Literature*, 42, 1009–1055.

Henrich, J., S. J. Heine, and A. Norenzayan (2010): "Most People Are Not WEIRD," *Nature*, 466, 29–29.

Holm, S. (1979): "A Simple Sequentially Rejective Multiple Test Procedure," *Scandinavian Journal of Statistics*, 6, 65–70.

Kershaw, C., D. E. Rast III, M. A. Hogg, and D. van Knippenberg (2021): "Battling Ingroup Bias with Effective Intergroup Leadership," *British Journal of Social Psychology*, 60, 765–785.

Kimbrough, E. O. and A. Vostroknutov (2016): "Norms Make Preferences Social," *Journal of the European Economic Association*, 14, 608–638.

Kline, P. and A. Santos (2012): "A Score Based Approach to Wild Bootstrap Inference," *Journal of Econometric Methods*, 1, 23–41.

Kosfeld, M. and D. Rustagi (2015): "Leader Punishment and Cooperation in Groups: Experimental Field Evidence from Commons Management in Ethiopia," *American Economic Review*, 105, 747–783.

Krupka, E. L. and R. A. Weber (2013): "Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?" *Journal of the European Economic Association*, 11, 495–524.

Lang, A. L., M. Omena, M. Y. Abdad, and R. L. Ford (2020): "Climate Change in Papua New Guinea: Impact on Disease Dynamics," *Papua New Guinea Medical Journal*, 58, 1–10.

Levati, M. V., M. Sutter, and E. van der Heijden (2007): "Leading by Example in a Public Goods Experiment with Heterogeneity and Incomplete Information," *Journal of Conflict Resolution*, 51, 793–818.

Lipset, S. M. (1959): *Political Man: The Social Bases of Politics*, Garden City, NJ: Doubleday.

Lipset, S. M. and S. Rokkan (1967): "Cleavage Structures, Party Systems, and Voter Alignments," in *Party Systems and Voter Alignments: Cross-National Perspectives*, New York: The Free Press, 1–64.

Morauta, L. (1974): *Beyond the Village. Local Politics in Madang, Papua New Guinea*, Routledge/Taylor & Francis Group.

Moxnes, E. and E. van der Heijden (2003): "The Effect of Leadership in a Public Bad Experiment," *Journal of Conflict Resolution*, 47, 773–795.

Nanau, G. (2011): "The Wantok System as a Socio-Economic and Political Network in Melanesia," *OMNES: The Journal of Multicultural Society*, 2, 31–55.

Oedzes, J. J., G. S. Van der Vegt, F. A. Rink, and F. Walter (2019): "On the Origins of Informal Hierarchy: The Interactive Role of Formal Leadership and Task Complexity," *Journal of Organizational Behavior*, 40, 311–324.

Ostrom, E. (2009): "A General Framework for Analyzing Sustainability of Social-Ecological Systems," *Science*, 325, 419–422.

——— (2010): "Analyzing Collective Action," *Agricultural Economics*, 41, 155–166.

Paolini, S., M. Gibbs, B. Sales, D. Anderson, and K. McIntyre (2024): "Negativity Bias in Intergroup Contact: Meta-analytical Evidence That Bad Is Stronger than Good, Especially When People Have the Opportunity and Motivation to Opt out of Contact," *Psychological Bulletin*, 150, 921–964.

Potters, J., M. Sefton, and L. Vesterlund (2007): "Leading-by-Example and Signaling in Voluntary Contribution Games: An Experimental Study," *Economic Theory*, 33, 169–182.

Regan, A. J. (2010): "Bougainville: Beyond Survival," *Cultural Survival*, https://www.culturalsurvival.org/publications/cultural–survival–quarterly/bougainville–beyond–survival (Accessed 03–14–2025).

Reilly, B. (2008): "Ethnic Conflict in Papua New Guinea," *Asia Pacific Viewpoint*, 49, 12–22.

Ross, L., D. Greene, and P. House (1977): "The "False Consensus Effect": An Egocentric Bias in Social Perception and Attribution Processes," *Journal of Experimental Social Psychology*, 13, 279–301.

Samek, A. S. and R. M. Sheremeta (2014): "Recognizing Contributors: An Experiment on Public Goods," *Experimental Economics*, 17, 673–690.

Samuelson, C. D. and S. T. Allison (1994): "Cognitive Factors Affecting the Use of Social Decision Heuristics in Resource-Sharing Tasks," *Organizational Behavior and Human Decision Processes*, 58, 1–27.

Scheepers, D. and N. Ellemers (2019): "Social Identity Theory," in *Social Psychology in Action*, ed. by K. Sassenberg and M. L. Vliek, Cham: Springer International Publishing, 129–143.

Schram, R. (2015): "A Society Divided: Death, Personhood, and Christianity in Auhelawa, Papua New Guinea," *HAU: Journal of Ethnographic Theory*, 5, 317–337.

Šidák, Z. (1967): "Rectangular Confidence Regions for the Means of Multivariate Normal Distributions," *Journal of the American Statistical Association*, 62, 626–633.

Sillanpää, M., D. A. Friess, C. D. Heatubun, S. M. Cragg, F. Alei, R. Bhargava, Wahyudi, Hendri, J. D. Kalor, Purwanto, C. Marlessy, R. P. Yudha, F. Sidik, D. Murdiyarso, and M. Lupascu (2024): "Mangrove Management Practices, Policies, and Impacts in New Guinea," *Biological Conservation*, 296, 110697.

Tajfel, H. (1978): "Social Categorization, Social Identity and Social Comparisons," in *Differentiation between Social Groups*, London, UK: Academic Press, 61–76.

Tajfel, H., M. G. Billig, R. P. Bundy, and C. Flament (1971): "Social Categorization and Intergroup Behaviour," *European Journal of Social Psychology*, 1, 149–178.

The World Bank (2025): "The World Bank In Papua New Guinea," https://www.worldbank.org/en/country/png/overview (Accessed 03–15–2025).

Turner, J. C. (1975): "Social Comparison and Social Identity: Some Prospects for Intergroup Behaviour," *European Journal of Social Psychology*, 5, 1–34.

Turner, J. C., M. A. Hogg, P. J. Oakes, S. D. Reicher, and M. S. Wetherell (1987): *Rediscovering the Social Group: A Self-Categorization Theory*, Rediscovering the Social Group: A Self-Categorization Theory, Cambridge, MA, US: Basil Blackwell.

Uğurlar, P., A. R. Dorrough, O. Isler, and O. Yilmaz (2025): "Shared Group Memberships Mitigate Intergroup Bias in Cooperation," *Social Psychological and Personality Science*, 16, 214–223.

Van Assche, J., H. Swart, K. Schmid, K. Dhont, A. Al Ramiah, O. Christ, M. Kauff, S. Rothmann, M. Savelkoul, N. Tausch, R. Wölfer, S. Zahreddine, M. Saleem, and M. Hewstone (2023): "Intergroup Contact Is Reliably Associated with Reduced Prejudice, Even in the Face of Group Threat and Discrimination," *The American Psychologist*, 78, 761–774.

Vousdoukas, M. I., P. Athanasiou, A. Giardino, L. Mentaschi, A. Stocchino, R. E. Kopp, P. Menéndez, M. W. Beck, R. Ranasinghe, and L. Feyen (2023): "Small Island Developing States under Threat by Rising Seas Even in a 1.5 °C Warming World," *Nature Sustainability*, 1–13.

Wadey, M., S. Brown, R. J. Nicholls, and I. Haigh (2017): "Coastal Flooding in the Maldives: An Assessment of Historic Events and Their Implications," *Natural Hazards*, 89, 131–159.

Weisel, O. and R. Böhm (2015): ""Ingroup Love" and "Outgroup Hate" in Intergroup Conflict between Natural Groups," *Journal of Experimental Social Psychology*, 60, 110–120.

Wells, S., C. Ravilious, and E. Corcoran (2006): *In the Front Line : Shoreline Protection and Other Ecosystem Services from Mangroves and Coral Reefs*, UNEP-WCMC.

Wiessner, P. (2020): "The Role of Third Parties in Norm Enforcement in Customary Courts among the Enga of Papua New Guinea," *Proceedings of the National Academy of Sciences*, 117, 32320–32328.

Wilkinson-Ryan, T. (2008): "The Sucker Norm," *U of Penn, Inst for Law & Econ Research Paper No. 09-05*.

Yu, J. and M. G. Kocher (2023): "Leading by Example in a Public Goods Experiment with Benefit Heterogeneity," *Social Choice and Welfare*, 61, 685–712.

# Online Appendix

Under Watchful Eyes: Observation by In-Group Leaders Promotes Intergroup Cooperation in Natural Groups

Juliane Koch & Jonas Pilgaard Kaiser

## A Changes from Preregistration

In this section, we describe the few changes we made from the preregistration to the paper:

- Number of participants: In the preregistration, the target sample size was intentionally left vague because – as stated in the preregistration – the feasible number of participants "depends on timing and budget", and the exact number of participants would therefore be "finally decided during the field phase". The preregistration thus lists potential sample sizes, from 240 subjects in the minimum case that participants could only be recruited in four villages to 300 or 360 participants if data could be collected in more villages. At the end, it was practically feasible to recruit a total of 402 participants across six villages, and we decided to do so to maximize statistical power. We planned to collect data in all six villages before the data collection started, and we did not look at the data during the data collection phase.

- Hypotheses: The wording of the hypotheses has been altered to make the connection to the experimental design and the statistical tests clearer. For example, the pre-registered Hypothesis 1 was as follows:
  "Hypothesis 1: Contribution levels are lower when cooperating with people in the out-group compared to the baseline (cooperation with in-group member without observation)."
  In the paper, the same hypothesis has been rewritten for clarity:
  "More participants cooperate in InPrivate than in OutPrivate."

- Statistical tests: The main analysis differs slightly from the pre-analysis plan by relying on logit regressions, chi-squared tests, and Fisher's exact tests instead of the Mann-Whitney U-test, which was preregistered but is less suitable for testing differences with binary outcomes. The pre-analysis plan also mentions treatment comparisons conditional on the participants' beliefs, and we discuss these results in Section 6.1 rather than in the main analysis.

# B Additional Study Setting Description

## B.1 Climate, Natural Resources, and Community-Based Adaptation

In this section, we provide further details about the impact of and adaptation to climate change in Papua New Guinea (PNG) in general and Bougainville in particular. Similar to other Pacific island states, PNG is heavily affected by climate change and its resulting natural hazards (Lang et al., 2020; Vousdoukas et al., 2023). Towards the year 2090, the Pacific Climate Change Science Program (PCCSP) expects that the climate in PNG will be marked by (i) increasing temperatures, (ii) an increase in the incidence of very hot days, (iii) changing rainfall patterns, (iv) an increase in the incidence of extreme rainfall days, and (v) less frequent but more intense tropical cyclones (Power et al., 2011). Extreme temperatures and frequent climate disasters such as cyclones, droughts, and flooding endanger vital harvests, especially for coastal villages (Wadey et al., 2017; Bourke, 2018). Recent calculations have shown that PNG has one of the highest 'Expected Annual Number of People Exposed' (EAPE) to coastal flooding (Vousdoukas et al., 2023). Sea level rise occurs twice as quickly in PNG compared to the global average with a pace of up to 10 mm/year (Nunn, 2013). In Bougainville, the share of coastal communities is high, and climate change has made the relocation of entire communities common practice (Bronen, 2014; Luetz and Havea, 2018), although the coastal villagers are often strongly reluctant to relocate (Davies, 2002).

Despite the country's vulnerability to climate change, PNG is richly endowed with natural resources which provide natural protection against various natural hazards. For instance, mangrove trees and coral reefs are of utmost importance for coastal communities of that region (Warner, 2000). With their ability to act as natural barriers against storm surges and coastal erosion, mangroves offer invaluable protection against the impacts of coastal flooding and rising sea levels (Alongi, 2002). In a similar way, coral reefs serve as natural breakwaters, reducing the vulnerability of coastal areas stemming from the wave energy of tsunamis and cyclones (Wells et al., 2006). Thus, conserving marine resources is key for providing local protection, as they reduce the impact of waves and storm surges (Wells et al., 2006). They further contribute to coastal resilience, as they generate soil accumulation and thus stabilize against erosion (Alongi, 2002).

All six villages in this study share the same characteristics of being a coastal community with a high abundance of mangrove trees. Yet, natural resources give rise to a critical social dilemma that all villages in this study face: On the one hand, exploiting natural resources such as mangrove trees can be a major source of income because its timber is used as building material and fire wood and commonly sold on local markets, and it is therefore financially

beneficial to harvest mangrove trees. On the other hand, its preservation is key for protecting coastal communities against natural hazards, such as coastal flooding or soil erosion, and this protection benefits everyone. Thus, preserving these natural resources requires cooperation across the villages that share the same coastline which is the case with the villages in our sample.

The contribution to 'community funds' is common practice in the daily village life in Bougainville. Such community funds are relevant when villagers contribute money, advice or time for common projects, such as jointly repairing parts of the community's church, or other public spaces. We therefore decided to use a similar concept in the experimental game in order to facilitate the understanding of the different strategies when deciding whether to cooperate. In order to further investigate biases on different group levels, we mimic a cooperation setting that not only involves cooperation of people *within* but also *across* villages. For this reason, we enlarge the well-known concept of the 'community fund' to a 'district climate fund'. To do so, we use the topic of mangrove protection which represents a relevant topic for the villages that share a coastline abundant in mangrove trees, as villagers are aware that the coastline protection requires cooperation across villages.

One of the features of using common funds for community-based adaptation is that it engages the people who are vulnerable to climate change themselves, thereby improving the chances of finding solutions that are contextually relevant and sustainable (Kearney et al., 2007; Nelson and Agrawal, 2008). The importance of community-based adaptation seems particularly relevant for a setting like Bougainville, where the trust in politicians on a higher level is extremely low. In our sample, we asked the villagers whether they trust (i) their national politicians, (ii) the politicians in Bougainville, (iii) their district politicians, (iv) the Ward politicians (comprising various communities in the region), and (v) the community leaders (binary question). For the national, Bougainville, and district politicians, only between 7 and 12 percent respond 'mostly yes' to whether they trust their politicians. In contrast, almost half of the participants (46 percent) respond 'mostly yes' to whether they trust their community leaders. Trust, shared norms, and social networks that facilitate cooperation are central for successful cooperation (Ostrom, 1990; Pretty and Ward, 2001). Yet, cooperation is often more difficult to sustain with out-groups, and most community-based adaptation initiatives to date are small stand-alone activities that operate on a local scale (Forsyth, 2013). This is, however, insufficient for dealing with climate impacts that transcend geographical boundaries and affect neighboring communities simultaneously.

## B.2 Sample Description

The main sample of the current experiment comprises 60-70 adults from each of the six villages, yielding a total sample of 402 villagers. On average, participants in this study are around 36 years old, and 52 percent are female. The average participant has completed around 8 years of schooling and report having a weekly income of 72 Kina, to which non-financial earnings through harvests, etc. are usually added. For this reason, the weekly income variable only provides a proxy for the participants' socio-economic status (we report a series of robustness checks to this effect in Section D). Other indicators of socio-economic status include the amount of money charged to the phone and the frequency of internet usage. Yet, given the sparse electric infrastructure in the study villages, many participants (79 percent) never use the internet, and 50 percent of the sample does not charge any money on their phone. We do see, however, that 20 percent of the sample uses the internet at least once a week, and 17 percent have more than 100 Kina on their phones. We also find that 72 percent of the villagers currently live in the village they were born in.

Bougainvillians are predominantly Christians and the weekly church visits are usually perceived as an important element of the villagers' social life. Thus, in our sample 77 percent of villagers report to attend church once a week, while 20 percent visit the church more often than that. Apart from religious encounters that represent an important component of Bougainvillian societies, community engagement is perceived as highly important. On average 89 percent report to engage in community activities on a regular basis, where "activities" imply that the villagers provide time, money, advice or a combination of these. This importance of community engagement corroborates our results from the vignettes presented in Appendix E.1, showing that free-riding is not socially appropriate.

Tables B.1 and B.2 break these summary statistics down by treatment, showing in general no major treatment differences, which is to be expected given the randomization of participants into treatments.

As described in the previous section, the island of Bougainville is heavily influenced by climate change. It is thus not surprising that on average 97 percent of our sample reports to believe in climate change. This has led many people to engage in measures that protect them against the consequences of climate change: On average, 72 percent of the sample reports to have engaged in some kind of climate adaptation measure before. Approximately 20 percent of the sample reports to not feel safe from environmental hazards in their coastal village, whereas 74 percent of the participants report that they have feared environmental hazards or environmental changes for food security reasons. Apart from food shortage, flooding is the second most-mentioned fear that causes participants to not feel safe from environmental

Table B.1: Demographics by Treatment

| | InPrivate | OutPrivate | InObserve | OutObserve | Total |
|---|---|---|---|---|---|
| Age | 34.82 | 37.63 | 38.71 | 34.57 | 36.24 |
| | (13.79) | (14.76) | (13.28) | (13.02) | (13.84) |
| Years of Schooling | 8.08 | 7.90 | 8.03 | 8.20 | 8.05 |
| | (2.21) | (2.73) | (2.17) | (2.27) | (2.36) |
| Weekly Income | 77.10 | 56.18 | 79.11 | 77.80 | 71.99 |
| | (135.51) | (80.29) | (143.89) | (219.17) | (149.99) |

*Notes:* The table shows mean values for each treatment and for the total sample, and it reports standard errors in parentheses. Weekly Income is measured in local currency Kina (1 EUR = 3.9 Kina).

hazards in their villages.

We also elicited self-reported knowledge, experience and attitudes of the participants with respect to environmental hazards influenced by climate change. In particular, we asked participants about eight environmental hazards: Floods, tsunamis, earthquakes, sea levels rise, droughts, erosion, change in weather patterns, and food shortage. For each environmental hazard, the participant indicated (i) whether they had heard about the hazard, (ii) whether they had directly experienced the hazard, and (iii) to what extent they think that the environmental hazard is important for Bougainville. Throughout, we find that hazards of food shortage, flooding, and drought take central stage, whereas earthquakes and erosions seem to matter less to the participants.

On average, the participants have heard about 84 percent of the eight environmental hazards. The most well-known are food shortage (98 percent) and floods (91 percent), whereas the least well-known are changes in weather patterns (73 percent) and erosion (74 percent). On average, participants have experienced 39 percent of the hazards.[1] Most participants report to have experienced food shortage (66 percent), drought (55 percent), and floods (52 percent), whereas the environmental hazards that the fewest participants report to have experienced are earthquakes (14 percent) and erosion (16 percent).

Participants are more likely to rate all hazards as "Somewhat relevant" than irrelevant or very relevant. The hazards that are rated "Very relevant" by the most participants are food shortage (26 percent), floods (23 percent), and drought (19 percent). In contrast, the hazards less likely to be rated as "Very relevant" include earthquakes (4 percent), weather change (8

[1]For the question of whether participants have been harmed by any of the environmental hazards, we do not ask about the change in weather patterns, as it is a more general notion that may be difficult for participants to confidently answer yes or no to.

Table B.2: Sample Characteristics by Treatment

| | InPrivate | OutPrivate | InObserve | OutObserve | Total |
|---|---|---|---|---|---|
| **Gender** | | | | | |
| Male | 50% | 50% | 39% | 49% | 48% |
| Female | 50% | 50% | 61% | 51% | 52% |
| **Frequency of Internet Usage** | | | | | |
| Never | 75% | 77% | 86% | 80% | 79% |
| Once a Month | 2% | 1% | 1% | 0% | 1% |
| Once a Week | 14% | 16% | 10% | 13% | 14% |
| Several Times a Week | 4% | 4% | 1% | 3% | 3% |
| Every day | 4% | 2% | 1% | 3% | 3% |
| **Money on Phone** | | | | | |
| 0 | 45% | 50% | 58% | 49% | 50% |
| K5 | 7% | 3% | 4% | 14% | 7% |
| K10 | 6% | 11% | 6% | 7% | 7% |
| K30 | 13% | 17% | 8% | 11% | 13% |
| K100 | 8% | 6% | 4% | 5% | 6% |
| +K100 | 21% | 13% | 19% | 13% | 17% |
| **Frequency of Church Attendance** | | | | | |
| Never | 2% | 3% | 0% | 1% | 2% |
| Once a Year | 2% | 0% | 0% | 1% | 1% |
| Once a Month | 0% | 2% | 0% | 1% | 1% |
| Once a Week | 76% | 75% | 80% | 78% | 77% |
| Several Times a Week | 15% | 9% | 7% | 8% | 10% |
| Every day | 6% | 12% | 13% | 11% | 10% |
| **Community Engagement** | | | | | |
| No | 8% | 13% | 8% | 14% | 11% |
| Yes | 92% | 87% | 92% | 86% | 89% |
| **Born in Village** | | | | | |
| No | 28% | 32% | 32% | 20% | 28% |
| Yes | 72% | 68% | 68% | 80% | 72% |

percent), and erosion (10 percent).

Table B.3 breaks the summary statistics across all hazards down by treatment, showing again no major treatment differences. This provides further support to the successful randomization of participants into treatments.

In Appendix E.2, we show that neither perceived safety from environmental hazards, engagement in climate adaptation measures, fear for food security, climate knowledge, climate attitudes, nor climate experience interacts with the treatments in the experiment. This suggests that although the framing was useful for helping participants understand the experiment, it did not influence the treatment effects.

Table B.3: Responses to Climate Questions by Treatment

| | InPrivate | OutPrivate | InObserve | OutObserve | Total |
|---|---|---|---|---|---|
| Believe in Climate Change | 0.97 | 0.98 | 0.94 | 0.97 | 0.97 |
| Engaged in Climate Adaptation | 0.70 | 0.75 | 0.75 | 0.69 | 0.72 |
| Safety Perception | 0.85 | 0.79 | 0.79 | 0.71 | 0.79 |
| Fear for Food Security | 0.64 | 0.78 | 0.81 | 0.79 | 0.74 |
| Climate Knowledge | 0.83 | 0.83 | 0.86 | 0.86 | 0.84 |
| Climate Attitudes | 0.37 | 0.39 | 0.38 | 0.42 | 0.39 |
| Climate Experience | 0.38 | 0.37 | 0.40 | 0.39 | 0.39 |

*Notes:* Believe in Climate Change, Engagement in Climate Adaptation, Perceived Safety from Disasters, and Fear for Food Security are all binary questions. Climate Knowledge is the share of natural hazards that the participant reported to have heard about. Climate Attitudes is the average relevance score given by the participant to each of the natural hazards (standardized between 0 and 1). Climate Experience is the share of the natural hazards that the participant has experienced.

# C Power Analysis

To examine the statistical power of our analysis with a sample of 402 participants, we conducted power simulations in Stata (version 18). To do this, we simulated data using the rbinomial command, and we used logit regressions for the analysis as in Section 5. For the power simulations, we decide to be agnostic about the potential explanatory power of the control variables that we use in our analysis, implying that we do not factor in the efficiency gain from including the control variables. For this reason, we consider the power estimates to be conservative.

For the power simulations, we assume as a baseline that the share of participants cooperating is 0.37, following the mega study on Prisoner's Dilemma games by Mengel (2018). We took this as the level of cooperation in the InPrivate treatment and examined our power to detect increases in cooperation following from observation and decreases in cooperation following from interacting with members of the out-group in private. Apart from the number of participants (see Appendix A), we do not update other parameters in our power analysis to

fit our actual data to avoid issues of ex-post power calculations (Hoenig and Heisey, 2001).

We expect a power of .8 to detect decreases in cooperation (following from out-group interaction in private, cf. H1) of approximately 16 percentage points. In addition, we expect a power of .8 to detect increases in cooperation (following from observation, cf. H2) of approximately 19 percentage points.

Calculating the power to detect the interaction effect is complicated by the fact that the power depends not only on the interaction but also the baseline effects of the group condition in private and the observation condition when interacting with in-groups. Computing the power with respect to the interaction effect is further complicated by the fact that the interaction may be both positive and negative, cf. H3. We examine the power to detect the interaction effect for all combinations of group effects of $\{0, .1, .2, .3\}$ and observation effects of $\{0, .05, .1, .15, .2, .25, .3\}$. Across all 21 combinations of different possible effects, we find that we have approximately .8 power to detect negative interaction effects of 15-27 percentage points and positive interaction effects of 12-26 percentage points.

# D Robustness

## D.1 Linear Probability Model

In the following, we examine the robustness of the results presented in Section 5 to using a linear probability model (LPM) instead of a logit regression. Looking first at H1, we find that the effect is robust to using LPM and statistically significant for all levels of controls ($p < .001$, cf. Table D.4). Using all control variables, we find that participants who decide in private on average become 25 percentage points less likely to contribute to the climate fund when interacting with a member of another village ($p < .001$).

Table D.4: H1: In-Group Bias and Contribution Levels, LPM

| | (1) | (2) | (3) |
|---|---|---|---|
| Out-group | -0.29*** | -0.28*** | -0.25*** |
| | (0.06) | (0.06) | (0.06) |
| Demographic Controls | No | Yes | Yes |
| Attitudinal Controls | No | No | Yes |
| Observations | 234 | 220 | 185 |

*Notes:* Linear Probability Model with contribution (binary) as the dependent variable. The sample comprises only participants who were not observed. The demographic controls are age, gender, years of schooling, weekly income, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Heteroskedasticity-robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Looking next at H2, we note that the result is also robust to using an LPM instead of logit regressions. Specifically, the effect is marginally significant without controls ($p = .066$), and it becomes significant when adding demographic controls ($p = .039$) and highly significant when also including attitudinal controls ($p = .009$), cf. Table D.5.

Table D.5: H2: Observation and Contribution Levels, LPM

| | (1) | (2) | (3) |
|---|---|---|---|
| Observation | 0.13* | 0.15** | 0.22*** |
| | (0.07) | (0.07) | (0.08) |
| Demographic Controls | No | Yes | Yes |
| Attitudinal Controls | No | No | Yes |
| Observations | 200 | 187 | 155 |

*Notes:* Linear Probability Model with contribution (binary) as the dependent variable. The sample comprises only participants who interacted with a member of their own village. The demographic controls are age, gender, years of schooling, weekly income, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Heteroskedasticity-robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Finally, we consider H3. First, we notice in a joint regression including the interaction between observation and group matching that participants are significantly less likely to con-

tribute when they are matched with out-groups (24 percentage points, $p < .001$), and that they are significantly more likely to contribute when they are matched with in-groups (23 percentage points, $p = .005$). The interaction is significant such that observation is even more important for out-groups (31 percentage points, $p = .004$), such that there is no effect of group matching when people are observed ($p = .400$). The interaction is statistically significant also when not including controls or when including only demographic controls ($p < .001$), cf. Table D.6.

Table D.6: H3: Interaction Effect of In-Group Bias and Observation on Contributions, LPM

| | (1) | (2) | (3) |
|---|---|---|---|
| Out-group | -1.45*** | -1.49*** | -0.24*** |
| | (0.31) | (0.33) | (0.06) |
| Observation | 0.54* | 0.64** | 0.23*** |
| | (0.30) | (0.31) | (0.08) |
| Observation × Out-group | 1.92*** | 1.85*** | 0.31*** |
| | (0.45) | (0.48) | (0.11) |
| Demographic Controls | No | Yes | Yes |
| Attitudinal Controls | No | No | Yes |
| Observations | 402 | 379 | 315 |

*Notes:* Linear Probability Model with contribution (binary) as the dependent variable. The demographic controls are age, gender, years of schooling, weekly income, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Heteroskedasticity-robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## D.2 Sample Restrictions Based on Control Questions

In the following, we examine the robustness of the results presented in Section 5 to excluding participants who make many mistakes when answering the control questions. As explained in Section 3, participants were asked a series of control questions to ensure their understanding of the experimental design. If the participants made two mistakes when answering the control questions, the experimenter explained again the rules of the experiment. Yet, because of time constraints in some sessions, not all eight control questions were asked to all participants. In this section, we show that our results are robust to various sample restrictions based on the control questions, suggesting that our results are not influenced by misunderstanding of the experimental instructions nor the fact that not all participants received all control questions.

We first examine the results concerning H1. As seen in Table D.7, the treatment effect is robust to and is statistically significant when (i) excluding the few participants who did not answer any control questions (all $p's < .001$), (ii) excluding participants who did not answer all control questions (all $p's < .001$), (iii) excluding participants who made more than four mistakes in total (all $p's < .009$), or (iv) excluding participants who did not answer four control questions correctly (all $p's < .001$).

Table D.7: H1: In-Group Bias and Contribution Levels, Restrictions From Control Questions

| | (1) | (2) | (3) |
|---|---|---|---|
| **Exclude participants without control questions** | | | |
| Out-group | -0.31*** | -0.29*** | -0.26*** |
| | (0.06) | (0.06) | (0.06) |
| Observations | 234 | 220 | 185 |
| **Exclude participants who did not answer all control questions** | | | |
| Out-group | -0.31*** | -0.29*** | -0.25*** |
| | (0.06) | (0.06) | (0.07) |
| Observations | 203 | 200 | 168 |
| **Exclude participants who made more than four mistakes in total** | | | |
| Out-group | -0.25*** | -0.24*** | -0.20*** |
| | (0.06) | (0.07) | (0.07) |
| Observations | 195 | 181 | 152 |
| **Exclude participants who did not answer four control questions correctly** | | | |
| Out-group | -0.29*** | -0.27*** | -0.24*** |
| | (0.06) | (0.06) | (0.06) |
| Observations | 204 | 201 | 171 |
| Demographic Controls | No | Yes | Yes |
| Attitudinal Controls | No | No | Yes |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The sample comprises only participants who were not observed. The demographic controls are age, gender, years of schooling, weekly income, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

We next examine the results concerning H2. As in the analysis in Section 5, the treatment effect is marginally statistically significant without controls, and it becomes statistically

significant when including all controls. As seen in Table D.8, the treatment effect is at least marginally statistically significant when (i) excluding the few participants who did not answer any control questions (all $p's < .092$), (ii) excluding participants who did not answer all control questions (all $p's < .087$), (iii) excluding participants who made more than four mistakes in total (all $p's < .077$), or (iv) excluding participants who did not answer four control questions correctly (all $p's < .075$).

Table D.8: H2: Observation and Contribution Levels, Restrictions From Control Questions

| | (1) | (2) | (3) |
|---|---|---|---|
| **Exclude participants without control questions** | | | |
| Observation | 0.13* | 0.14* | 0.20** |
| | (0.08) | (0.07) | (0.08) |
| Observations | 185 | 181 | 150 |
| **Exclude participants who did not answer all control questions** | | | |
| Observation | 0.14* | 0.15* | 0.20** |
| | (0.08) | (0.08) | (0.09) |
| Observations | 166 | 163 | 135 |
| **Exclude participants who made more than four mistakes in total** | | | |
| Observation | 0.14* | 0.16* | 0.23** |
| | (0.08) | (0.08) | (0.09) |
| Observations | 170 | 157 | 127 |
| **Exclude participants who did not answer four control questions correctly** | | | |
| Observation | 0.14* | 0.16* | 0.20** |
| | (0.08) | (0.08) | (0.09) |
| Observations | 170 | 166 | 137 |
| Demographic Controls | No | Yes | Yes |
| Attitudinal Controls | No | No | Yes |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The sample comprises only participants who were not observed. The demographic controls are age, gender, years of schooling, weekly income, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Finally, we examine the results concerning H3. As seen in Table D.9, the interaction effect is statistically significant when (i) excluding the few participants who did not answer any

control questions (all $p's < .002$), (ii) excluding participants who did not answer all control questions (all $p's < .002$), (iii) excluding participants who made more than four mistakes in total (all $p's < .018$), or (iv) excluding participants who did not answer four control questions correctly (all $p's < .004$).

Table D.9: H3: Interaction Effect of In-Group Bias and Observation on Contributions, Restrictions From Control Questions

| | (1) | (2) | (3) |
|---|---|---|---|
| **Exclude participants without control questions** | | | |
| Out-group | -0.29*** | -0.29*** | -0.25*** |
| | (0.05) | (0.05) | (0.06) |
| Observation | 0.11 | 0.13* | 0.19** |
| | (0.07) | (0.07) | (0.08) |
| Observation × Out-group | 0.40*** | 0.38*** | 0.35*** |
| | (0.07) | (0.08) | (0.10) |
| Observations | 375 | 368 | 305 |
| **Exclude participants who did not answer all control questions** | | | |
| Out-group | -0.29*** | -0.28*** | -0.25*** |
| | (0.05) | (0.05) | (0.06) |
| Observation | 0.12* | 0.14* | 0.20** |
| | (0.08) | (0.08) | (0.08) |
| Observation × Out-group | 0.40*** | 0.38*** | 0.36*** |
| | (0.08) | (0.08) | (0.10) |
| Observations | 344 | 338 | 280 |
| **Exclude participants who made more than four mistakes in total** | | | |
| Out-group | -0.24*** | -0.23*** | -0.20*** |
| | (0.06) | (0.06) | (0.07) |
| Observation | 0.13* | 0.16* | 0.23** |
| | (0.08) | (0.08) | (0.09) |
| Observation × Out-group | 0.36*** | 0.33*** | 0.29** |
| | (0.09) | (0.09) | (0.12) |
| Observations | 347 | 324 | 269 |
| **Exclude participants who did not answer four control questions correctly** | | | |
| Out-group | -0.28*** | -0.27*** | -0.23*** |
| | (0.05) | (0.06) | (0.06) |
| Observation | 0.13* | 0.15** | 0.21** |
| | (0.07) | (0.08) | (0.09) |
| Observation × Out-group | 0.39*** | 0.36*** | 0.34*** |
| | (0.08) | (0.09) | (0.11) |
| Observations | 355 | 348 | 290 |
| Demographic Controls | No | Yes | Yes |
| Attitudinal Controls | No | No | Yes |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The demographic controls are age, gender, years of schooling, weekly income, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

### D.3 Restricting to All Control Variables Elicited

In the analysis presented in Section 5, all specifications include all participants who responded to the control variables used for the respective specification. While this likely maximizes the power within each specification, as it uses the data from more participants, it also implies that specifications differ both in what variables are used as well as in what participants are included. In the following, we examine the robustness of the results to instead using throughout only data from the participants who answered the demographic and attitudinal variables.

Looking first at H1, we find that the effect is robust in all specifications to including only the participants who answered all demographic and attitudinal variables ($p < .001$, cf. Table D.10).

The results for H2 are also robust to including only the participants who provide answers for all demographic and attitudinal control variables. Specifically, the effect is statistically significant in all specifications (all $p's < .009$, cf. Table D.10).

Finally, we consider H3. We again find that our results are robust to including only the participants who answer all demographic and attitudinal questions. In particular, the interaction is significant such that observation is even more important for out-groups (all $p's < .002$, cf. Table D.10), and there is no effect of group matching when people are observed ($p = .433$).

Table D.10: Restricting to Participants Who Answered All Controls

| | (1) | (2) | (3) |
|---|---|---|---|
| **H1: In-Group Bias and Contribution Levels** | | | |
| Out-group | -0.22*** | -0.23*** | -0.24*** |
| | (0.06) | (0.06) | (0.06) |
| Observations | 185 | 185 | 185 |
| **H2: Observation and Contribution Levels** | | | |
| Observation | 0.24*** | 0.23*** | 0.22*** |
| | (0.08) | (0.08) | (0.08) |
| Observations | 155 | 155 | 155 |
| **H3: Interaction Effect of In-Group Bias and Observation on Contributions** | | | |
| Out-group | -0.23*** | -0.24*** | -0.24*** |
| | (0.06) | (0.06) | (0.06) |
| Observation | 0.22*** | 0.21*** | 0.21*** |
| | (0.08) | (0.08) | (0.08) |
| Observation × Out-group | 0.33*** | 0.33*** | 0.34*** |
| | (0.10) | (0.10) | (0.10) |
| Observations | 315 | 315 | 315 |
| Demographic Controls | No | Yes | Yes |
| Attitudinal Controls | No | No | Yes |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The demographic controls are age, gender, years of schooling, weekly income, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

## D.4 Excluding Participants Flagged by Experimenters

When conducting the experiment, the experimenters flagged a total of 10 participants (2.5 percent of the sample) where it was unclear whether the session worked as intended, for example, if the village leader in the observed condition seemed absent-minded, or if the participant tried to cooperate by contributing another amount than the K7 stated in the instructions. As we show below, the results are robust to excluding these 10 participants.

Looking first at H1, we find that the effect is robust in all specifications to excluding the 10 flagged participants ($p < .001$, cf. Table D.11).

The results for H2 are also robust to excluding the 10 flagged participants. Specifically,

the effect is at least marginally statistically significant in all specifications (all $p's < .068$, cf. Table D.11) and statistically significant when including control variables.

Finally, we consider H3. When excluding the 10 flagged participants, we find that the interaction is significant such that observation is even more important for out-groups (all $p's < .001$, cf. Table D.11), and there is no effect of group matching when people are observed ($p = .219$).

Table D.11: Restricting to Participants Who Were Not Flagged by Experimenters

| | (1) | (2) | (3) |
|---|---|---|---|
| **H1: In-Group Bias and Contribution Levels** | | | |
| Out-group | -0.30*** | -0.29*** | -0.24*** |
| | (0.06) | (0.06) | (0.06) |
| Observations | 231 | 217 | 182 |
| **H2: Observation and Contribution Levels** | | | |
| Observation | 0.13* | 0.15** | 0.22*** |
| | (0.07) | (0.07) | (0.08) |
| Observations | 198 | 185 | 153 |
| **H3: Interaction Effect of In-Group Bias and Observation on Contributions** | | | |
| Out-group | -0.28*** | -0.28*** | -0.24*** |
| | (0.05) | (0.05) | (0.06) |
| Observation | 0.12* | 0.13* | 0.20** |
| | (0.07) | (0.07) | (0.08) |
| Observation × Out-group | 0.44*** | 0.42*** | 0.40*** |
| | (0.07) | (0.07) | (0.09) |
| Observations | 392 | 369 | 307 |
| Demographic Controls | No | Yes | Yes |
| Attitudinal Controls | No | No | Yes |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The demographic controls are age, gender, years of schooling, weekly income, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

### D.5 Excluding Village Leaders From the Sample

Before the village leaders observed the participants in the observed conditions, they were randomly assigned to either the InPrivate or the OutPrivate treatment and made a decision in the experiment. Because village leaders are thus also villagers taking part in the experiment, the analysis presented in Section 5 includes the data from eight village leaders. Yet, as we discuss in Section E.3, there may be different expectations concerning the village leaders than other villagers. In this section, we show that our results are robust to excluding the eight village leaders.

Looking first at H1, we find that the effect is robust in all specifications to excluding the village leaders ($p < .001$, cf. Table D.12).

The results for H2 are also robust to excluding the village leaders. Specifically, the effect is at least marginally statistically significant in all specifications (all $p's < .085$, cf. Table D.12) and statistically significant when including control variables.

Finally, we consider H3. When excluding the village leaders, we find that the interaction is significant such that observation is even more important for out-groups (all $p's < .001$, cf. Table D.12), and there is no effect of group matching when people are observed ($p = .433$).

Table D.12: Excluding Village Leaders From the Sample

| | (1) | (2) | (3) |
|---|---|---|---|
| **H1: In-Group Bias and Contribution Levels** | | | |
| Out-group | -0.30*** | -0.29*** | -0.24*** |
| | (0.06) | (0.06) | (0.06) |
| Observations | 226 | 216 | 185 |
| **H2: Observation and Contribution Levels** | | | |
| Observation | 0.13* | 0.15** | 0.22*** |
| | (0.07) | (0.07) | (0.08) |
| Observations | 196 | 185 | 155 |
| **H3: Interaction Effect of In-Group Bias and Observation on Contributions** | | | |
| Out-group | -0.29*** | -0.28*** | -0.24*** |
| | (0.05) | (0.05) | (0.06) |
| Observation | 0.11* | 0.14* | 0.21*** |
| | (0.07) | (0.07) | (0.08) |
| Observation × Out-group | 0.40*** | 0.38*** | 0.34*** |
| | (0.07) | (0.08) | (0.10) |
| Observations | 394 | 375 | 315 |
| Demographic Controls | No | Yes | Yes |
| Attitudinal Controls | No | No | Yes |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The demographic controls are age, gender, years of schooling, weekly income, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

## D.6 Income Proxies

As explained in Section 3, we seek to control for the income of the participants, and we control for the participants' self-reported weekly incomes in the main analysis presented in Section 5. Yet, because the majority of people does not earn money but live in a self-sufficient manner, often using barter, questions about monetary income do not provide sufficient information about how materially well-off participants are. We therefore also ask participants about their money-recharging values on their cell phones, and we ask participants about how often they use the internet. Such questions about cell phones and internet use also have drawbacks, however, as many villagers do not have cell phones or may have little to no internet reception

in the village. We show in this section that our results are robust to the different proxies for the income of participants.

In the following, we test the hypotheses using each of the three income proxies separately and jointly, using internet frequency and money-recharging value on phone as discrete variables. In doing so, we note that the money-recharging value correlates significantly with internet frequency (Spearman's $\rho = .483$, $p < .001$) and weekly income (Spearman's $\rho = .139$, $p < .006$), whereas internet frequency and weekly income are uncorrelated (Spearman's $\rho = .003$, $p < .961$).

Examining first H1, we find that the treatment effect is robust and statistically significant across all specifications (all $p's < .001$, cf. Table D.13).

Table D.13: H1: In-Group Bias and Contribution Levels, Different Income Proxies

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Out-group | -0.28*** | -0.24*** | -0.29*** | -0.25*** | -0.30*** | -0.25*** | -0.28*** | -0.25*** |
| | (0.06) | (0.06) | (0.06) | (0.06) | (0.06) | (0.06) | (0.06) | (0.06) |
| Weekly Income | 0.00 | 0.00 | | | | | 0.00 | 0.00 |
| | (0.00) | (0.00) | | | | | (0.00) | (0.00) |
| Phone: K5 | | | 0.10 | 0.18 | | | 0.08 | 0.23 |
| | | | (0.17) | (0.21) | | | (0.17) | (0.24) |
| Phone: K10 | | | -0.02 | 0.06 | | | 0.00 | 0.18 |
| | | | (0.11) | (0.12) | | | (0.12) | (0.13) |
| Phone: K30 | | | -0.08 | -0.07 | | | -0.08 | -0.03 |
| | | | (0.10) | (0.10) | | | (0.10) | (0.10) |
| Phone: K100 | | | -0.07 | -0.05 | | | -0.07 | -0.01 |
| | | | (0.11) | (0.12) | | | (0.12) | (0.13) |
| Phone: +K100 | | | -0.03 | 0.01 | | | 0.00 | 0.08 |
| | | | (0.09) | (0.11) | | | (0.10) | (0.13) |
| I: Once a Week | | | | | -0.10 | -0.10 | -0.07 | -0.11 |
| | | | | | (0.09) | (0.09) | (0.10) | (0.09) |
| I: Several Times a Week | | | | | 0.03 | -0.17 | 0.04 | -0.22*** |
| | | | | | (0.15) | (0.11) | (0.16) | (0.08) |
| I: Every day | | | | | -0.14 | -0.17* | -0.14 | -0.19** |
| | | | | | (0.13) | (0.09) | (0.13) | (0.09) |
| Demographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Attitudinal Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 220 | 185 | 220 | 185 | 212 | 177 | 212 | 177 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The sample comprises only participants who were not observed. The discrete "I:" variable describes the frequency of internet use with "Never" as the baseline category and "Once a Month" omitted because of too few observations. The demographic controls are age, gender, years of schooling, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Looking next at H2, we find that the treatment effect is at least marginally statistically significant across all specifications. As seen in Table D.14, all specifications that do not include attitudinal control variables yield at least marginally statistically significant effects (all $p's < .094$), and all specifications that include also attitudinal controls yield statistically significant results (all $p's < .026$).

Table D.14: H2: Observation and Contribution Levels, Different Income Proxies

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Observation | 0.15** | 0.22*** | 0.13* | 0.19** | 0.13* | 0.19** | 0.13* | 0.20** |
| | (0.07) | (0.08) | (0.07) | (0.08) | (0.08) | (0.08) | (0.07) | (0.08) |
| Weekly Income | -0.00 | -0.00* | | | | | -0.00 | -0.00 |
| | (0.00) | (0.00) | | | | | (0.00) | (0.00) |
| Phone: K5 | | | -0.17 | -0.30 | | | -0.06 | -0.03 |
| | | | (0.17) | (0.19) | | | (0.17) | (0.22) |
| Phone: K10 | | | -0.03 | 0.10 | | | 0.02 | 0.20 |
| | | | (0.16) | (0.18) | | | (0.17) | (0.18) |
| Phone: K30 | | | -0.21* | -0.27** | | | -0.12 | -0.16 |
| | | | (0.12) | (0.13) | | | (0.13) | (0.14) |
| Phone: K100 | | | -0.02 | -0.03 | | | 0.05 | 0.07 |
| | | | (0.15) | (0.15) | | | (0.14) | (0.15) |
| Phone: +K100 | | | -0.07 | -0.05 | | | 0.05 | 0.10 |
| | | | (0.09) | (0.10) | | | (0.10) | (0.11) |
| I: Once a Month | | | | | -0.21 | -0.13 | -0.23 | -0.24 |
| | | | | | (0.23) | (0.21) | (0.23) | (0.21) |
| I: Once a Week | | | | | -0.23** | -0.29*** | -0.24** | -0.31*** |
| | | | | | (0.10) | (0.09) | (0.10) | (0.09) |
| I: Several Times a Week | | | | | -0.13 | -0.14 | -0.13 | -0.26 |
| | | | | | (0.23) | (0.21) | (0.24) | (0.16) |
| I: Every day | | | | | -0.23 | -0.21 | -0.28* | -0.29** |
| | | | | | (0.16) | (0.14) | (0.16) | (0.14) |
| Demographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Attitudinal Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 187 | 155 | 188 | 156 | 187 | 155 | 186 | 154 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The sample comprises only participants who interacted with a member of their own village. The discrete "I:" variable describes the frequency of internet use with "Never" as the baseline category. The demographic controls are age, gender, years of schooling, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Finally, looking at H3, we find that our results are robust to the different proxies for income, as the interaction is positive and statistically significant across all specifications, implying that observation is even more important for out-groups (all $p's < .002$, cf. Table D.16). Across all specifications, we again see that there is no effect of group matching when people are observed (all $p's > .260$).

Table D.15: H3: Interaction Effect of In-Group Bias and Observation on Contributions, Different Income Proxies

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Out-group | -0.28*** | -0.24*** | -0.28*** | -0.23*** | -0.28*** | -0.24*** | -0.28*** | -0.23*** |
| | (0.05) | (0.06) | (0.05) | (0.06) | (0.05) | (0.06) | (0.05) | (0.06) |
| Observation | 0.14** | 0.21*** | 0.13* | 0.19** | 0.12* | 0.18** | 0.13* | 0.20** |
| | (0.07) | (0.08) | (0.07) | (0.08) | (0.07) | (0.08) | (0.07) | (0.08) |
| Observation × Out-group | 0.38*** | 0.34*** | 0.38*** | 0.35*** | 0.39*** | 0.37*** | 0.38*** | 0.35*** |
| | (0.08) | (0.10) | (0.08) | (0.10) | (0.08) | (0.10) | (0.08) | (0.10) |
| Weekly Income | -0.00 | -0.00 | | | | | -0.00 | -0.00 |
| | (0.00) | (0.00) | | | | | (0.00) | (0.00) |
| Phone: K5 | | | -0.02 | -0.03 | | | 0.02 | 0.05 |
| | | | (0.11) | (0.13) | | | (0.11) | (0.14) |
| Phone: K10 | | | -0.09 | -0.01 | | | -0.06 | 0.07 |
| | | | (0.10) | (0.11) | | | (0.10) | (0.11) |
| Phone: K30 | | | -0.07 | -0.07 | | | -0.03 | -0.00 |
| | | | (0.08) | (0.08) | | | (0.08) | (0.08) |
| Phone: K100 | | | -0.03 | -0.01 | | | -0.01 | 0.02 |
| | | | (0.10) | (0.11) | | | (0.10) | (0.11) |
| Phone: +K100 | | | -0.01 | 0.04 | | | 0.05 | 0.12 |
| | | | (0.07) | (0.08) | | | (0.08) | (0.08) |
| I: Once a Month | | | | | -0.19 | -0.18 | -0.20 | -0.24* |
| | | | | | (0.18) | (0.15) | (0.17) | (0.14) |
| I: Once a Week | | | | | -0.08 | -0.10 | -0.09 | -0.13 |
| | | | | | (0.08) | (0.08) | (0.08) | (0.08) |
| I: Several Times a Week | | | | | -0.03 | -0.18 | -0.02 | -0.22* |
| | | | | | (0.14) | (0.13) | (0.15) | (0.13) |
| I: Every day | | | | | -0.25* | -0.24* | -0.27** | -0.29** |
| | | | | | (0.13) | (0.13) | (0.12) | (0.11) |
| Demographic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Attitudinal Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 379 | 315 | 380 | 316 | 373 | 309 | 372 | 308 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The sample comprises only participants who were not observed. The discrete "I:" variable describes the frequency of internet use with "Never" as the baseline category. The demographic controls are age, gender, years of schooling, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## D.7 Village Fixed Effects

In the following, we examine the robustness of the results presented in Section 5 to including village fixed effects in all specifications. Looking first at H1, we find that the effect is robust to including village fixed effects and statistically significant for all levels of controls ($p < .001$,

cf. Table D.16). Using all control variables, we find that participants who decide in private on average become 25 percentage points less likely to cooperate when interacting with a member of another village than with a member of their own village ($p < .001$).

The results for H2 are also robust to including village fixed effects. Specifically, the effect is statistically significant in all specifications (all $p's < .032$, cf. Table D.16).

Finally, we consider H3. We again find that our results are robust to including village fixed effects, as the interaction is significant such that observation is even more important for out-groups (32 percentage points, $p = .003$, cf. Table D.16), implying that there is no effect of group matching when people are observed ($p = .519$).

### D.8 Clustering

Clustering occurs when error terms are correlated within groups but are independent across groups, and it poses problems for statistical inference: the correlation of the error terms often leads to standard errors that are biased downwards, thereby causing an over-rejection of the null hypotheses (Cameron and Miller, 2015). Such groups could e.g. be villages in the current study, where one might worry that correlation in behavior between villagers from the same village could influence the results of our inference. In the main specifications reported in Section 5, we estimate models with standard errors that are robust to heteroscedasticity but not to clustering. The reason for this is that participants were randomized into treatments within villages, ensuring that the treatment variable is uncorrelated with possible clusters. As Cameron and Miller (2015) note (p. 334): "If a key regressor is randomly assigned within clusters, or is as good as randomly assigned, then the within-cluster correlation of the regressor is likely to be zero. Thus, there is no need to cluster standard errors even if the model's errors are clustered." To test whether this holds in our setting, we compute the intraclass correlation (ICC) for the villages using the ANOVA estimator (loneway in Stata 18). The ICC measures the proportion of residual variance attributable to differences between villages. Across all hypotheses and specifications, we find negligible ICCs: For H1, we find ICCs of $\rho = .003$, $\rho = .007$, and $\rho = .015$. For H2, they are $\rho = .024$, $\rho = .017$, and $\rho = .017$. And for H3, they are $\rho < .001$, $\rho = .019$, and $\rho = .021$. All 9 ICCs have confidence intervals that include zero and upper bounds that never exceed .090. This indicates that village-level clustering explains little to no residual variance, and this corroborates the approach taken in Section 5, where we do not adjust the heteroscedasticity-robust standard errors to clustering.

Still, we examine in this section the robustness of our results for using cluster-robust inference. First, we examine the results using cluster-robust standard errors (CRSE). Across the different levels of control, we find that the treatment effects are statistically significant

Table D.16: Including Village Fixed Effects

| | (1) | (2) | (3) |
|---|---|---|---|
| **H1: In-Group Bias and Contribution Levels, Village Fixed Effects** | | | |
| Out-group | -0.29*** | -0.28*** | -0.25*** |
| | (0.06) | (0.06) | (0.06) |
| Observations | 234 | 220 | 185 |
| **H2: Observation and Contribution Levels, Village Fixed Effects** | | | |
| Observation | 0.15** | 0.18** | 0.24*** |
| | (0.07) | (0.07) | (0.08) |
| Observations | 200 | 187 | 155 |
| **H3: Interaction Effect of In-Group Bias and Observation on Contributions** | | | |
| Out-group | -0.27*** | -0.28*** | -0.23*** |
| | (0.05) | (0.05) | (0.06) |
| Observation | 0.12* | 0.14** | 0.22*** |
| | (0.07) | (0.07) | (0.08) |
| Observation × Out-group | 0.39*** | 0.37*** | 0.32*** |
| | (0.07) | (0.08) | (0.10) |
| Observations | 402 | 379 | 315 |
| Demographic Controls | No | Yes | Yes |
| Attitudinal Controls | No | No | Yes |
| Village Fixed Effects | Yes | Yes | Yes |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The demographic controls are age, gender, years of schooling, weekly income, and community engagement. The attitudinal controls are perceptions of altruism, perceptions of fairness, trust, and risk preferences. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

when using CRSE for both H1 (all $p's < .001$), H2 ($p = .060$, $p = .032$, and $p = .009$), and H3 ($p = .002$, $p = .013$, and $p = .008$).

One problem with using CRSE is that it tends to underestimate standard errors when the number of clusters is small, leading to inflated Type I error rates (Cameron and Miller, 2015). As we have only six villages in our sample, we therefore also examine the robustness of our results to using the score wild cluster bootstrap, which is a method tailored to conducting inference with few clusters after maximum likelihood estimation (e.g. our primary specification relying on logit models, Kline and Santos, 2012). To estimate the models with only six clusters,

we use the weights proposed by Webb (2023), which are recommended for conducting wild cluster bootstrap for fewer than 10 clusters (Cameron and Miller, 2015). Again, we find across the different levels of control that the treatment effects are statistically significant when using score wild cluster bootstrap for both H1 ($p = .009$, $p = .005$, and $p = .002$), H2 ($p = .079$, $p = .076$, and $p = .015$), and H3 ($p = .032$, $p = .066$, and $p = .091$).

# E Discussion

## E.1 Vignettes

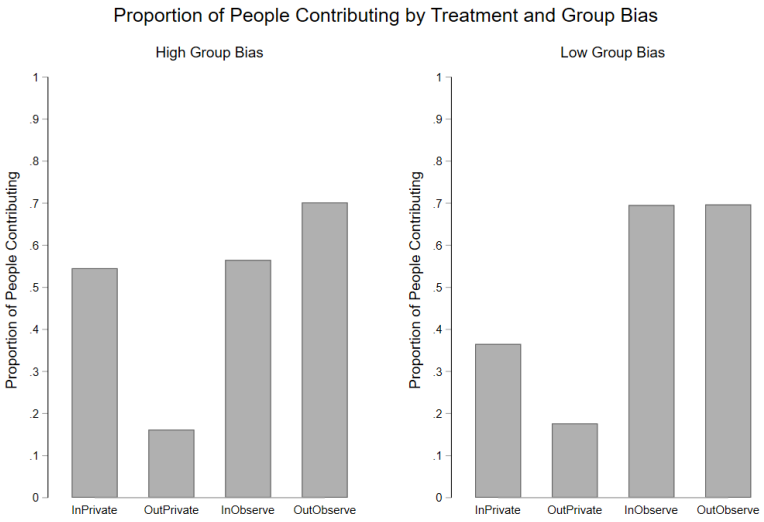
### E.1.1 Group Bias

In the last part of the post-experimental survey, we further examine the motivations and attitudes of the villagers through three vignettes. The first vignette describes a person, Mr. Fred, who receives money from an NGO to plant 20 trees to protect villages against flooding. Participants learn that all villages need trees to protect against climate change. We then elicit attitudes towards Mr. Fred's tree planting behavior, where participants can choose between the following options: A) Plant all 20 trees in his own village; B) Plant 10 trees in his own village and leave the other 10 for the two closest villages; C) Plant 4 trees in his own village and leave the other 16 for other villages in the northern area of Bougainville; D) Plant 1 tree in his own village and leave the other 19 for other villages over the whole of Bougainville. Thus, the response options show different levels of own-village favoritism, with A indicating the most in-group bias and D the least.

In the survey, we find large variation in what villagers think Mr. Fred should do: Specifically, 34 percent choose Option A, 28 percent choose Option B, 18 percent choose Option C, and 20 percent choose Option D.

The vignette does not include any strategic considerations, such as repeated interaction or reciprocity from the other villages. Consequently, the vignette elicits in-group favoritism that is more closely related to preferences for one's in-group than beliefs about the behavior of another individual. Thus, if the group variation in the experiment, analyzed in Section 5.1, measures in-group bias, and if behavior in the experiment is driven by preferences for the outcomes of the in-group and out-group, respectively, we would expect that answers to this vignette correlate with the treatment effect for H1. This is precisely what we find when we separate the sample into below-median in-group bias (answers C and D) and above-median in-group bias (answers A and B): We find a statistically significant effect between InPrivate and OutPrivate among the most biased individuals (contributions: 55 percent vs. 16 percent; $p = .001$, logit), while the effect on the least biased individuals is only marginally significant

(contributions: 37 percent vs. 18 percent; $p = .057$, logit).




Notes: The figure presents the average contribution level for each treatment, dividing participants into those with above-median in-group bias (answers C and D in the first vignette) and blow-median in-group bias (answers A and B in the first vignette).

### E.1.2 Free-Riding Is Undesirable

As explained in Section 4, we expected that observation by the village leader would result in more contributions if cooperation is perceived as socially desirable. In Section 5.3, we asked about beliefs concerning the views of the village leader and the village in general. This analysis showed that villagers on average think that others prefer that they cooperate, regardless of whether they interact with a villager from the same or another village.

Our vignettes corroborate that villagers view contribution as fair. Specifically, one vignette elicits the villagers' fairness views towards free-riding on others' contributions. The vignette describes two individuals, Steven and Peter, who live in a community that has introduced a climate community fund. The participants learn that Steve has contributed with time and money into the community fund, whereas Peter has not contributed anything. Then, a natural disaster happens. Thanks to the climate community fund, both Steven's and Peter's harvests are saved. The participants are then asked to state on a four-point Likert scale how fair they think it is that Peter did not contribute to the climate community fund. We find that almost all participants (95 percent) find it either "Not fair at all" or "Rather not fair" that Peter did not contribute, with the remaining 5 percent saying that it is "Rather fair". Thus, there is an almost unanimous attitude that not contributing to a climate fund that one benefits from is

unfair.

To put the almost unanimous agreement on the unfairness of free-riding into perspective, we can compare this to the fairness views on luck egalitarianism. One vignette presents participants with two women, Thelma and Iris, who invest the same time and effort into two gardens of the same size. The participants learn that after a month of work, a big storm destroys nearly all of Iris's harvest, while Thelma's harvest survives entirely, resulting in her earning a lot of money from harvest market sales. When asked whether it is fair that Thelma earns much more money than Iris, 22 percent report this to be "Very fair" or "Rather fair", while 78 percent think this is "Not fair at all" or "Rather not fair".

#### E.1.3 Punishment Attitudes

As described above, a vignette in the post-experimental survey explains a situation in which a person (Peter) does not contribute to a climate fund that he benefits from. While the above results deal with the fairness of Peter's behavior, further questions shed light on the participants' attitudes towards punishment. The participant learns that the community thinks about excluding Peter from the benefits of the climate fund if he continues not to contribute. The participants are then asked whether they find it fair that the community excludes Peter from the benefits in case a natural disaster happens. We find that although participants on average think it is unfair not to contribute to a community climate fund that everybody benefits from, they on average do not find it fair to punish such behavior. Specifically, 73 percent of the participants say that it is "Not fair at all" or "Rather not fair", with the remaining 27 percent say that it would be "Very fair" or "Rather fair".

Similarly, in the vignette describing how a person (Thelma) earned more due to not being harmed by a natural disaster, further questions ask about how gossip or punishment by a third person would influence the third person's reputation. Here, we find that 86 percent (2 percent) of participants say that gossiping would lead to a worse (better) reputation, and 97 percent (0 percent) of participants say that punishment would lead to a worse (better) reputation for the person who gossips or punishes, respectively.

### E.2 External Validity of the Lab-in-the-Field Experiment

As explained in Section 3, the experimental instructions applied the context of climate change to enhance the participants' understanding of the Prisoner's Dilemma game. One concern with using a framed experiment is that the framing may interact with the treatment effects, reducing the external validity of the study's conclusions. Yet, as we show in this section, there are no significant interaction effects between the treatment effects and the participants'

knowledge about environmental hazards, experiences with environmental hazards, attitudes towards climate change, perceived safety from climate disasters, or fear for food security due to environmental hazards. This indicates that the framing of the current experiment did not influence our reported treatment effects, thereby supporting the generalizability of the study's findings.

In the post-experimental survey, we elicited self-reported knowledge, experience and attitudes of the participants with respect to environmental hazards influenced by climate change. In particular, we asked participants about eight environmental hazards: Floods, tsunamis, earthquakes, sea levels rise, droughts, erosion, change in weather patterns, and food shortage. For each environmental hazard, the participant indicated (i) whether they had heard about the hazard, (ii) whether they had directly experienced the hazard, and (iii) to what extent they think that the environmental hazard is important for Bougainville.

We examine first possible treatment interactions with self-reported knowledge about environmental hazards. We use a median split to indicate those who report to know the most environmental hazards, and we include all controls as in our preferred specification in Section 5. Doing so, we find no significant interaction between climate knowledge and treatments for H1 ($p = .256$), H2 ($p = .367$), nor H3 ($p = .222$).

Second, we examine possible treatment interactions with whether participants have experienced the different environmental hazards. Again, we use a median split to indicate those who report to have experienced the most environmental hazards. Again, we find no significant interactions with the treatments for H1 ($p = .547$), H2 ($p = .632$), nor H3 ($p = .841$).

Third, we examine possible interactions with attitudes towards environmental hazards. In particular, we aggregate across all environmental hazards how important participants find the hazards to be. We then use a median split to indicate the participants who on average rate environmental hazards to be the most important. Again, we find no interaction with the treatment effects for H1 ($p = .304$), H2 ($p = .941$), nor H3 ($p = .916$).

Fourth, we examine interactions with whether participants say that they believe they are safe from disasters (binary Yes/No). We again find no interaction with the treatment effects for H1 ($p = .952$), H2 ($p = .656$), nor H3 ($p = .221$).

Fifth, we examine interactions with whether participants have feared environmental hazards for food security reasons. We again find no interaction with the treatment effects for H1 ($p = .322$), H2 ($p = .857$), nor H3 ($p = .741$).

In sum, the treatment effects do not interact with beliefs, attitudes, or experiences with environmental hazards. This suggests that the treatment effects are not influenced by the climate framing used in the current experiment.

### E.3 Different Expectations Towards Village Leaders

The analyses in Sections 5.2 and 5.3 document that observation by village leaders induced more villagers to cooperate. In all treatments, this is related to participants believing that the village leader wants them to contribute, and that they would dislike it if the participant deviates from this expectation. Yet, the beliefs of the participants also reveal that the expectation about intergroup cooperation may not hold for the village leaders themselves. Combining responses from all treatments, approximately 70 percent believe that the village leader would cooperate when interacting with someone from their own village. Yet, only around 20 percent believe that the village leader would contribute if the other person is from another village. Both of these are statistically significantly different from 50 percent (binomial tests, $p's < .001$). This leads to a discrepancy in out-group interactions between what participants believe that the village leader wants them to do and what participants believe the village leader themselves would do: In OutPrivate and OutObserve, participants most often believe that the village leader wants them to contribute while the village leader would not themselves contribute (OutPrivate: 59 percent, McNemar's test: $p < .001$; OutObserve: 60 percent, McNemar's test: $p < .001$). In contrast, beliefs about the attitudes and behavior of the village leader mostly coincide for InPrivate and InObserve (InPrivate: 66 percent, McNemar's test: $p = .108$; InObserve: 66 percent, McNemar's test: $p = .275$).

A similar discrepancy between the villagers and the expected behavior of the village leaders is seen by comparing beliefs about in-group bias among male and female village leaders. While male villagers exhibit in-group bias to a greater extent than female villagers (cf. Section E.4.1), participants do not believe that the bias differs between male and female village leaders: Most participants believe that their village leader would act similarly regardless of sex in both situations (in-group: 82 percent, McNemar's test: $p = .901$; out-group: 85 percent, McNemar's test: $p = .500$). Thus, participants believe that both their male and female village leaders would behave differently towards members of their in-group and out-group (McNemar's test: $p's < .001$), and the difference in expected in-group bias is not significant (sign test, $p = .586$).

This indicates that there may exist different expectations concerning the behavior of the villagers and the behavior of village leaders. Previous research has also shown that there are different standards for leaders and other group members: Leaders are often given more leeway to deviate from group norms (Smith, 2020), and group members are more forgiving when in-group leaders misbehave compared to when other in-group members or out-group leaders misbehave (Abrams et al., 2013). We view it as an interesting avenue for future research to examine whether double standards for group leaders influence intergroup cooperation when (unlike in our setting) intergroup cooperation depends more on decisions made by leaders than

decisions made by individual group members.

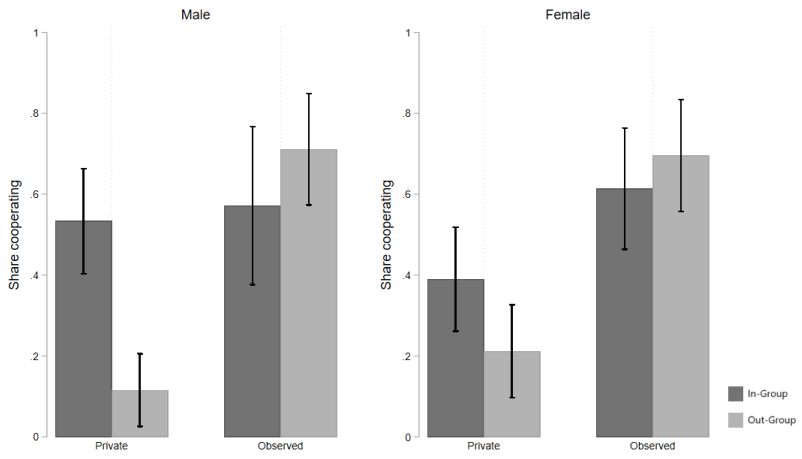
## E.4 Subgroup Analysis

In this section, we examine differences in behavior for different subgroups of our data. In particular, we look at possible demographics that may predict contribution in itself and which may interact with the treatment effects. Throughout, we use logit regressions with a similar estimation strategy to the one used in Section 5.

### E.4.1 Gender Differences

Previous research has documented gender differences in in-group bias and strategic signaling behavior. For example, Fershtman and Gneezy (2001) show in a Jewish society that ethnic discrimination is strong among men but non-existent among women, and Vugt et al. (2007) find that men respond more strongly than women to intergroup competition. At the same time, men and women often respond differently to observation: To signal formidability and toughness, for instance, men are more likely to initiate negotiations (Kugler et al., 2018), use sabotage to improve performance (Dato and Nieken, 2014), or decrease cooperation when acting in front of their peers (Charness and Rustichini, 2011). Consequently, it is possible that men and women behave differently in the current experiment, both in responsiveness to the group affiliation of the other player and to observation by the village leader.

Figure E.1: Contribution Behavior Across Treatments and Sex




Notes: This figure shows the share of participants contributing for each treatment. Error bars indicate 95 percent confidence intervals.

In this study, there is no overall significant difference between the general willingness to contribute for men and women (logit, $p = .640$). Yet, women are much less affected by the group affiliation of the other player: Whereas the contribution rate among men decreases by 35 percentage points when interacting with someone from another village (logit, $p < .001$), the decrease among women is a statistically insignificant 13 percentage points (logit, $p = .165$). Consequently, there is less scope for observation to have a bias-reducing effect for women. That is, while observation substantially reduces out-group bias among men (APE: 40 percent, logit: $p < .001$), the effect is somewhat less pronounced for women (APE: 28 percent, logit: $p = .021$).

One possible reason for the more pronounced in-group bias among men could be the gender differences in interactions with members of other villages in this field setting. Such interactions take place only rarely due to the seclusion of villages and lack of terrestrial infrastructure (see Section 2). Yet, when they meet, men typically interact with members of other villages in competitive settings. In most self-sufficient families, for example, men are responsible for fishing, and here they compete with fishers from other villages and have no close interactions. In contrast, women have closer contact in a more cooperative setting when meeting women from other villages. For instance, villages that are geographically closer share some facilities such as schools or medical facilities, and women are usually responsible for bringing and picking up school children. Contact with members of other villages may influence how people view their out-group (cf. the contact hypothesis, Allport, 1954; Paluck et al., 2019). According to the contact hypothesis, when the environment for such encounters is cooperative, contact may reduce in-group bias.[2] As the typical environment for contact with members of other villages is more cooperative for women, this could explain why women exhibit less in-group bias in the current study.

### E.4.2 Age

We first examine the direct correlation of age with villagers' inclination to contribute. We find no correlation between age and the probability that participants contribute, neither in the treatments separately ($APE \in [-.0033; .0042]$, all $p's > .287$) nor combined ($APE = .0005$, $p = .773$). The results are unchanged when adding controls.

Second, we examine whether age interacts with the treatment effects. Regardless of the level of controls, we find no significant interaction with age neither for H1 (all $p's > .151$), H2

[2]Specifically, Allport (1954) describe that appropriate conditions for intergroup contact to reduce prejudice is when (i) the two groups have an equal status (as opposed to differences in prestige or rank), (ii) the groups work towards common goals, (iii) the groups cooperate rather than compete to reach their goals, and (iv) an authority, law, or custom encourages interaction between the groups.

(all $p's > .357$), nor H3 (all $p's > .718$).

### E.4.3 Income

We examine the correlation between contribution and income using a median split, comparing individuals with above-median weekly income to individuals with below-median weekly income. Looking first at the direct correlation, we find no correlation between having above-median weekly income and contributing in the experiment, neither in the treatments separately ($APE \in [-.0574; .0377]$, all $p's > .624$) nor combined ($APE = .0122$, $p = .793$). The results are unchanged when adding controls.

Second, we examine whether the level of income interacts with the treatment effects. Regardless of the level of controls, we find no significant interaction with having above-median weekly income neither for H1 (all $p's > .900$), H2 (all $p's > .522$), nor H3 (all $p's > .760$).

### E.4.4 Education

We examine the correlation between contribution and education using a median split, comparing individuals with above-median education to individuals with below-median education. Looking first at the direct correlation with having above-median education, we find a positive correlation between having above-median education and contributing in the experiment in In-Private that just reaches statistical significance ($APE = .19$, $p = .044$), and this is marginally statistically significant when adding controls ($APE = .18$, $p = .075$). Yet, the correlation is not present in any of the other treatments, and the effect across all treatments is insignificant ($APE = .0661$, $p = .173$). We thus conclude that, overall, education does not predict participants' inclination to contribute in the experiment.

Second, we examine whether the level of education interacts with the treatment effects. Regardless of the level of controls, we find no significant interaction with having above-median education neither for H1 (all $p's > .857$), H2 (all $p's > .125$), nor H3 (all $p's > .519$).

### E.4.5 Birth Village

Lastly, we examine whether the inclination to contribute in the experiment is associated with whether the villager was born in another village. Looking first at whether having a different birth village predicts contributions, we find that whether a respondent lives in the same village as they were born does not predict whether one cooperates in the experiment, neither in the treatments separately ($APE \in [-.0972; .0175]$, all $p's > .426$) or combined ($APE = .0548$, $p = .288$). The results are unchanged when adding controls.

Second, we examine whether having a different birth village interacts with the treatment effects. Regardless of the level of controls, we find no significant interaction with having been born in another village neither for H1 (all $p's > .527$), H2 (all $p's > .672$), nor H3 (all $p's > .575$).

# F Additional Tables and Figures

## F.1 Main Hypotheses

Table F.17: H1: In-Group Bias and Contributions

| | (1) | (2) | (3) |
|---|---|---|---|
| Out-group | -0.29*** | -0.28*** | -0.24*** |
| | (0.06) | (0.06) | (0.06) |
| Age | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Female | | -0.03 | -0.03 |
| | | (0.06) | (0.06) |
| Years of Schooling | | 0.04*** | 0.03** |
| | | (0.01) | (0.01) |
| Weekly Income | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Community Engagement | | -0.13 | -0.14 |
| | | (0.10) | (0.10) |
| Altruism Perception | | | 0.00 |
| | | | (0.01) |
| Fairness Perception | | | -0.01 |
| | | | (0.01) |
| Trust, General | | | -0.04 |
| | | | (0.05) |
| Risk Aversion | | | -0.03** |
| | | | (0.01) |
| Baseline Cooperation Rate | .32 | .32 | .27 |
| Observations | 234 | 220 | 185 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The sample comprises only participants who were not observed. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table F.18: H2: Observation and Contribution Levels

| | (1) | (2) | (3) |
|---|---|---|---|
| Observation | 0.13* | 0.15** | 0.22*** |
| | (0.07) | (0.07) | (0.08) |
| Age | | 0.01** | 0.01* |
| | | (0.00) | (0.00) |
| Female | | -0.10 | -0.13 |
| | | (0.07) | (0.08) |
| Years of Schooling | | 0.04*** | 0.05*** |
| | | (0.02) | (0.02) |
| Weekly Income | | -0.00 | -0.00** |
| | | (0.00) | (0.00) |
| Community Engagement | | -0.11 | -0.13 |
| | | (0.12) | (0.12) |
| Altruism Perception | | | -0.00 |
| | | | (0.01) |
| Fairness Perception | | | -0.02 |
| | | | (0.02) |
| Trust, General | | | 0.03 |
| | | | (0.07) |
| Risk Aversion | | | -0.04** |
| | | | (0.02) |
| Baseline Cooperation Rate | .51 | .51 | .47 |
| Observations | 200 | 187 | 155 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The sample comprises only participants who interacted with a member of their own village. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Table F.19: H3: Interaction Effect of In-Group Bias and Observation on Contributions

| | (1) | (2) | (3) |
|---|---|---|---|
| Out-group | -0.28*** | -0.28*** | -0.24*** |
| | (0.05) | (0.05) | (0.06) |
| Observation | 0.12* | 0.14** | 0.21*** |
| | (0.07) | (0.07) | (0.08) |
| Observation × Out-group | 0.40*** | 0.38*** | 0.34*** |
| | (0.07) | (0.08) | (0.10) |
| Age | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Female | | -0.02 | -0.02 |
| | | (0.05) | (0.05) |
| Years of Schooling | | 0.03*** | 0.03** |
| | | (0.01) | (0.01) |
| Weekly Income | | -0.00 | -0.00 |
| | | (0.00) | (0.00) |
| Community Engagement | | -0.07 | -0.07 |
| | | (0.08) | (0.08) |
| Altruism Perception | | | 0.00 |
| | | | (0.01) |
| Fairness Perception | | | -0.00 |
| | | | (0.01) |
| Trust, General | | | -0.03 |
| | | | (0.04) |
| Risk Aversion | | | -0.02 |
| | | | (0.01) |
| Observations | 402 | 379 | 315 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Table F.20: Contributions and In-Group Bias, Observation

| | (1) | (2) | (3) |
|---|---|---|---|
| Out-group | 0.11 | 0.09 | 0.10 |
| | (0.07) | (0.08) | (0.08) |
| Age | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Female | | -0.01 | -0.01 |
| | | (0.08) | (0.09) |
| Years of Schooling | | 0.02 | 0.03 |
| | | (0.02) | (0.02) |
| Weekly Income | | -0.00 | -0.00 |
| | | (0.00) | (0.00) |
| Community Engagement | | -0.04 | -0.01 |
| | | (0.13) | (0.13) |
| Altruism Perception | | | -0.01 |
| | | | (0.02) |
| Fairness Perception | | | 0.01 |
| | | | (0.02) |
| Trust, General | | | 0.02 |
| | | | (0.07) |
| Risk Aversion | | | 0.00 |
| | | | (0.02) |
| Observations | 168 | 159 | 130 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The sample comprises only participants who were observed. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

## F.2 Discussion: Beliefs

Table F.21: Contributions and Beliefs About Other Player's Behavior

| | (1) | (2) | (3) |
|---|---|---|---|
| Expected Contribution | 0.16*** | 0.15*** | 0.15*** |
| | (0.05) | (0.05) | (0.05) |
| OutPrivate | -0.29*** | -0.28*** | -0.23*** |
| | (0.06) | (0.06) | (0.06) |
| InObserve | 0.13* | 0.15** | 0.22*** |
| | (0.07) | (0.07) | (0.08) |
| OutObserve | 0.26*** | 0.26*** | 0.32*** |
| | (0.06) | (0.07) | (0.07) |
| Age | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Female | | -0.02 | -0.02 |
| | | (0.05) | (0.05) |
| Years of Schooling | | 0.02** | 0.03** |
| | | (0.01) | (0.01) |
| Weekly Income | | -0.00 | -0.00 |
| | | (0.00) | (0.00) |
| Community Engagement | | -0.08 | -0.07 |
| | | (0.08) | (0.09) |
| Altruism Perception | | | 0.00 |
| | | | (0.01) |
| Fairness Perception | | | -0.00 |
| | | | (0.01) |
| Trust, General | | | -0.04 |
| | | | (0.04) |
| Risk Aversion | | | -0.02 |
| | | | (0.01) |
| Observations | 380 | 373 | 309 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Table F.22: Expected Contributions and In-Group Bias, No Observation (H1)

| | (1) | (2) | (3) |
|---|---|---|---|
| Out-group | -0.13* | -0.12* | -0.10 |
| | (0.07) | (0.07) | (0.07) |
| Age | | 0.00 | -0.00 |
| | | (0.00) | (0.00) |
| Female | | -0.03 | -0.04 |
| | | (0.07) | (0.07) |
| Years of Schooling | | 0.01 | 0.02 |
| | | (0.01) | (0.02) |
| Weekly Income | | -0.00 | -0.00 |
| | | (0.00) | (0.00) |
| Community Engagement | | 0.19* | 0.24** |
| | | (0.11) | (0.10) |
| Altruism Perception | | | -0.01 |
| | | | (0.01) |
| Fairness Perception | | | -0.00 |
| | | | (0.02) |
| Trust, General | | | 0.02 |
| | | | (0.06) |
| Risk Aversion | | | 0.01 |
| | | | (0.02) |
| Observations | 219 | 216 | 181 |

*Notes:* Logit regression with expected contribution of the other player (binary) as the dependent variable. The sample comprises only participants who were not observed. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Table F.23: Expected Contributions and Observation (H2)

| | (1) | (2) | (3) |
|---|---|---|---|
| Observation | -0.09 | -0.07 | -0.10 |
| | (0.08) | (0.08) | (0.08) |
| Age | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Female | | -0.06 | -0.08 |
| | | (0.07) | (0.08) |
| Years of Schooling | | 0.04** | 0.05*** |
| | | (0.02) | (0.02) |
| Weekly Income | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Community Engagement | | 0.00 | -0.00 |
| | | (0.14) | (0.14) |
| Altruism Perception | | | -0.01 |
| | | | (0.02) |
| Fairness Perception | | | -0.01 |
| | | | (0.02) |
| Trust, General | | | 0.13* |
| | | | (0.07) |
| Risk Aversion | | | -0.01 |
| | | | (0.02) |
| Observations | 186 | 182 | 150 |

*Notes:* Logit regression with expected contribution of the other player (binary) as the dependent variable. The sample comprises only participants who interacted with a member of their own village. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Table F.24: Expected Contributions and Observation/Group Bias (H3)

| | (1) | (2) | (3) |
|---|---|---|---|
| Out-group | -0.12* | -0.12* | -0.11 |
| | (0.07) | (0.07) | (0.07) |
| Observation | -0.09 | -0.08 | -0.09 |
| | (0.07) | (0.07) | (0.08) |
| Observation × Out-group | 0.02 | -0.00 | -0.01 |
| | (0.10) | (0.10) | (0.11) |
| Age | | 0.00 | -0.00 |
| | | (0.00) | (0.00) |
| Female | | -0.02 | 0.00 |
| | | (0.05) | (0.06) |
| Years of Schooling | | 0.02 | 0.02 |
| | | (0.01) | (0.01) |
| Weekly Income | | -0.00 | -0.00 |
| | | (0.00) | (0.00) |
| Community Engagement | | 0.05 | 0.06 |
| | | (0.08) | (0.09) |
| Altruism Perception | | | -0.01 |
| | | | (0.01) |
| Fairness Perception | | | -0.01 |
| | | | (0.01) |
| Trust, General | | | 0.06 |
| | | | (0.05) |
| Risk Aversion | | | 0.00 |
| | | | (0.01) |
| Observations | 380 | 373 | 309 |

*Notes:* Logit regression with expected contribution of the other player (binary) as the dependent variable. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table F.25: Contributions and In-Group Bias, No Observation, Including Beliefs (H1)

| | (1) | (2) | (3) |
|---|---|---|---|
| Out-group | -0.30*** | -0.28*** | -0.24*** |
| | (0.06) | (0.06) | (0.06) |
| Belief, Other Player Contributes | 0.09 | 0.08 | 0.08 |
| | (0.06) | (0.06) | (0.06) |
| Age | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Female | | -0.02 | -0.03 |
| | | (0.06) | (0.06) |
| Years of Schooling | | 0.03*** | 0.03** |
| | | (0.01) | (0.01) |
| Weekly Income | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Community Engagement | | -0.14 | -0.16 |
| | | (0.10) | (0.10) |
| Altruism Perception | | | 0.00 |
| | | | (0.01) |
| Fairness Perception | | | -0.01 |
| | | | (0.01) |
| Trust, General | | | -0.04 |
| | | | (0.06) |
| Risk Aversion | | | -0.03** |
| | | | (0.01) |
| Observations | 219 | 216 | 181 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The sample comprises only participants who were not observed. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Table F.26: Observation and Contribution Levels, In-Group, Including Beliefs (H2)

| | (1) | (2) | (3) |
|---|---|---|---|
| Observation | 0.13* | 0.15** | 0.22*** |
| | (0.07) | (0.07) | (0.08) |
| Belief, Other Player Contributes | 0.20*** | 0.17** | 0.15** |
| | (0.07) | (0.07) | (0.07) |
| Age | | 0.01* | 0.00 |
| | | (0.00) | (0.00) |
| Female | | -0.09 | -0.12 |
| | | (0.07) | (0.08) |
| Years of Schooling | | 0.04** | 0.04** |
| | | (0.02) | (0.02) |
| Weekly Income | | -0.00 | -0.00** |
| | | (0.00) | (0.00) |
| Community Engagement | | -0.10 | -0.12 |
| | | (0.12) | (0.12) |
| Altruism Perception | | | 0.00 |
| | | | (0.01) |
| Fairness Perception | | | -0.01 |
| | | | (0.01) |
| Trust, General | | | 0.00 |
| | | | (0.07) |
| Risk Aversion | | | -0.04** |
| | | | (0.02) |
| Observations | 186 | 182 | 150 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. The sample comprises only participants who interacted with a member of their own village. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Table F.27: Contributions and Observation/Group Bias, Including Beliefs (H3)

| | (1) | (2) | (3) |
|---|---|---|---|
| Out-group | -0.28*** | -0.27*** | -0.23*** |
| | (0.05) | (0.05) | (0.06) |
| Observation | 0.12* | 0.14* | 0.20*** |
| | (0.07) | (0.07) | (0.08) |
| Observation × Out-group | 0.41*** | 0.38*** | 0.35*** |
| | (0.07) | (0.08) | (0.10) |
| Belief, Other Player Contributes | 0.16*** | 0.14*** | 0.14*** |
| | (0.04) | (0.04) | (0.05) |
| Age | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Female | | -0.02 | -0.02 |
| | | (0.05) | (0.05) |
| Years of Schooling | | 0.02** | 0.03** |
| | | (0.01) | (0.01) |
| Weekly Income | | -0.00 | -0.00 |
| | | (0.00) | (0.00) |
| Community Engagement | | -0.08 | -0.07 |
| | | (0.08) | (0.09) |
| Altruism Perception | | | 0.00 |
| | | | (0.01) |
| Fairness Perception | | | -0.00 |
| | | | (0.01) |
| Trust, General | | | -0.04 |
| | | | (0.04) |
| Risk Aversion | | | -0.02 |
| | | | (0.01) |
| Observations | 380 | 373 | 309 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

## F.3 Discussion: Climate

Table F.28: Contributions and Climate Attitudes

| | (1) | (2) | (3) |
|---|---|---|---|
| Climate Attitudes | 0.19* | 0.18* | 0.23** |
| | (0.10) | (0.10) | (0.12) |
| OutPrivate | -0.30*** | -0.29*** | -0.24*** |
| | (0.06) | (0.06) | (0.06) |
| InObserve | 0.13* | 0.16** | 0.23*** |
| | (0.07) | (0.07) | (0.08) |
| OutObserve | 0.23*** | 0.23*** | 0.28*** |
| | (0.07) | (0.07) | (0.08) |
| Age | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Female | | -0.02 | -0.03 |
| | | (0.05) | (0.05) |
| Years of Schooling | | 0.03** | 0.03** |
| | | (0.01) | (0.01) |
| Weekly Income | | -0.00 | -0.00 |
| | | (0.00) | (0.00) |
| Community Engagement | | -0.06 | -0.06 |
| | | (0.08) | (0.08) |
| Altruism Perception | | | 0.00 |
| | | | (0.01) |
| Fairness Perception | | | 0.00 |
| | | | (0.01) |
| Trust, General | | | -0.04 |
| | | | (0.04) |
| Risk Aversion | | | -0.02 |
| | | | (0.01) |
| Observations | 385 | 378 | 315 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. Climate Attitudes is the average relevance score given by the participant to each of eight different natural hazards (standardized between 0 and 1). Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Table F.29: Contributions and Binary Climate Attitudes

| | (1) | (2) | (3) |
|---|---|---|---|
| High Climate Attitudes | 0.11** | 0.11** | 0.13*** |
| | (0.05) | (0.05) | (0.05) |
| OutPrivate | -0.30*** | | -0.24*** |
| | (0.06) | | (0.06) |
| InObserve | 0.14* | | 0.23*** |
| | (0.07) | | (0.08) |
| OutObserve | 0.24*** | | 0.30*** |
| | (0.07) | | (0.07) |
| Age | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Female | | 0.00 | -0.03 |
| | | (0.05) | (0.05) |
| Years of Schooling | | 0.03*** | 0.03** |
| | | (0.01) | (0.01) |
| Weekly Income | | 0.00 | -0.00 |
| | | (0.00) | (0.00) |
| Community Engagement | | -0.04 | -0.05 |
| | | (0.08) | (0.08) |
| Altruism Perception | | | 0.00 |
| | | | (0.01) |
| Fairness Perception | | | 0.00 |
| | | | (0.01) |
| Trust, General | | | -0.04 |
| | | | (0.04) |
| Risk Aversion | | | -0.01 |
| | | | (0.01) |
| Observations | 385 | 378 | 315 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. Whereas Climate Attitudes is the average relevance score given by the participant to each of eight different natural hazards (standardized between 0 and 1), High Climate Attitudes is a binary variable that takes on a value of 1 if the participant belongs to the 50 percent with the greatest average relevance score. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Table F.30: Contributions and Perceived Safety

| | (1) | (2) | (3) |
|---|---|---|---|
| Perceived Safety | -0.06 | -0.05 | -0.12* |
| | (0.06) | (0.06) | (0.07) |
| OutPrivate | -0.30*** | -0.30*** | -0.25*** |
| | (0.06) | (0.06) | (0.06) |
| InObserve | 0.13* | 0.14* | 0.21*** |
| | (0.07) | (0.07) | (0.08) |
| OutObserve | 0.22*** | 0.22*** | 0.27*** |
| | (0.07) | (0.07) | (0.08) |
| Age | | 0.00 | 0.00 |
| | | (0.00) | (0.00) |
| Female | | -0.01 | -0.01 |
| | | (0.05) | (0.05) |
| Years of Schooling | | 0.03** | 0.03** |
| | | (0.01) | (0.01) |
| Weekly Income | | -0.00 | -0.00 |
| | | (0.00) | (0.00) |
| Community Engagement | | -0.07 | -0.06 |
| | | (0.08) | (0.08) |
| Altruism Perception | | | 0.00 |
| | | | (0.01) |
| Fairness Perception | | | -0.01 |
| | | | (0.01) |
| Trust, General | | | -0.03 |
| | | | (0.04) |
| Risk Aversion | | | -0.02 |
| | | | (0.01) |
| Observations | 382 | 375 | 313 |

*Notes:* Logit regression with contribution (binary) as the dependent variable. Coefficients are average partial effects, heteroskedasticity-robust standard errors in parentheses.

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

Table F.31: Attitudes by Treatment

| | InPrivate | OutPrivate | InObserve | OutObserve | Total |
|---|---|---|---|---|---|
| Risk Aversion | 7.94 | 7.81 | 7.69 | 7.49 | 7.75 |
| | (2.03) | (2.18) | (2.39) | (2.29) | (2.20) |
| General Trust | 2.50 | 2.40 | 2.62 | 2.59 | 2.52 |
| | (0.54) | (0.60) | (0.58) | (0.59) | (0.58) |
| Perceived Altruism | 7.17 | 7.20 | 6.82 | 7.14 | 7.11 |
| | (2.47) | (2.36) | (2.42) | (2.33) | (2.39) |
| Perceived Fairness | 6.66 | 6.22 | 6.79 | 6.18 | 6.45 |
| | (2.41) | (2.58) | (2.34) | (2.66) | (2.51) |

*Notes:* The table shows mean values for each treatment and in the total sample, and it reports standard errors in parentheses. Risk Aversion, Perceived Alturism, and Perceived Fairness are all measured on scales from 1-10, General Trust is measured on a scale from 1-4.

# References


Abrams, D., G. Randsley de Moura, and G. A. Travaglino (2013): "A Double Standard When Group Members Behave Badly: Transgression Credit to Ingroup Leaders," *Journal of Personality and Social Psychology*, 105, 799–815.

Allport, G. W. (1954): *The Nature of Prejudice*, Cambridge, MA: Perseus Books.

Alongi, D. M. (2002): "Present State and Future of the World's Mangrove Forests," *Environmental conservation*, 29, 331–349.

Bourke, R. M. (2018): "Impact of Climate Change on Agriculture in Papua New Guinea," *Climate Change: Our Environment, Livelihoods, and Sustainabilit*, 35–50.

Bronen, R. (2014): "Community Relocations: The Arctic and South Pacific," in *Humanitarian Crises and Migration*, Routledge, 221–242.

Cameron, A. C. and D. L. Miller (2015): "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources*, 50, 317–372.

Charness, G. and A. Rustichini (2011): "Gender Differences in Cooperation with Group Membership," *Games and Economic Behavior*, 72, 77–85.

Dato, S. and P. Nieken (2014): "Gender Differences in Competition and Sabotage," *Journal of Economic Behavior & Organization*, 100, 64–80.

Davies, H. (2002): "Tsunamis and the Coastal Communities of Papua New Guinea," in *Natural Disasters and Cultural Change*, Routledge.

Fershtman, C. and U. Gneezy (2001): "Discrimination in a Segmented Society: An Experimental Approach*," *The Quarterly Journal of Economics*, 116, 351–377.

Forsyth, T. (2013): "Community-Based Adaptation: A Review of Past and Future Challenges," *WIREs Climate Change*, 4, 439–446.

Hoenig, J. M. and D. M. Heisey (2001): "The Abuse of Power," *The American Statistician*, 55, 19–24.

Kearney, J., F. Berkes, A. Charles, E. Pinkerton, and M. Wiber (2007): "The Role of Participatory Governance and Community-Based Management in Integrated Coastal and Ocean Management in Canada," *Coastal Management*.

Kline, P. and A. Santos (2012): "A Score Based Approach to Wild Bootstrap Inference," *Journal of Econometric Methods*, 1, 23–41.

Kugler, K. G., J. A. M. Reif, T. Kaschner, and F. C. Brodbeck (2018): "Gender Differences in the Initiation of Negotiations: A Meta-Analysis," *Psychological Bulletin*, 144, 198–222.

Lang, A. L., M. Omena, M. Y. Abdad, and R. L. Ford (2020): "Climate Change in Papua New Guinea: Impact on Disease Dynamics," *Papua New Guinea Medical Journal*, 58, 1–10.

Luetz, J. and P. H. Havea (2018): ""We're Not Refugees, We'll Stay Here until We Die!"—Climate Change Adaptation and Migration Experiences Gathered from the Tulun and Nissan Atolls of Bougainville, Papua New Guinea," *Climate change impacts and adaptation strategies for coastal communities*, 3–29.

Mengel, F. (2018): "Risk and Temptation: A Meta-study on Prisoner's Dilemma Games," *The Economic Journal*, 128, 3182–3209.

Nelson, F. and A. Agrawal (2008): "Patronage or Participation? Community-based Natural Resource Management Reform in Sub-Saharan Africa," *Development and change*, 39, 557–585.

Nunn, P. D. (2013): "The End of the Pacific? Effects of Sea Level Rise on Pacific Island Livelihoods," *Singapore Journal of Tropical Geography*, 34, 143–171.

Ostrom, E. (1990): *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge university press.

Paluck, E. L., S. A. Green, and D. P. Green (2019): "The Contact Hypothesis Re-Evaluated," *Behavioural Public Policy*, 3, 129–158.

Power, S., A. Schiller, G. Cambers, D. Jones, and K. Hennessy (2011): "The Pacific Climate Change Science Program," *Bulletin of the American Meteorological Society*, 92, 1409–1411.

Pretty, J. and H. Ward (2001): "Social Capital and the Environment," *World development*, 29, 209–227.

Smith, J. R. (2020): "Group Norms," in *Oxford Research Encyclopedia of Psychology*, Oxford University Press.

Vousdoukas, M. I., P. Athanasiou, A. Giardino, L. Mentaschi, A. Stocchino, R. E. Kopp, P. Menéndez, M. W. Beck, R. Ranasinghe, and L. Feyen (2023): "Small Island Developing States under Threat by Rising Seas Even in a 1.5 °C Warming World," *Nature Sustainability*, 1–13.

Vugt, M. V., D. D. Cremer, and D. P. Janssen (2007): "Gender Differences in Cooperation and Competition: The Male-Warrior Hypothesis," *Psychological Science*, 18, 19–23.

Wadey, M., S. Brown, R. J. Nicholls, and I. Haigh (2017): "Coastal Flooding in the Maldives: An Assessment of Historic Events and Their Implications," *Natural Hazards*, 89, 131–159.

Warner, M. (2000): "Conflict Management in Community-Based Natural Resource Projects: Experiences from Fiji and Papua New Guinea," *Overseas Development Institute*.

Webb, M. D. (2023): "Reworking Wild Bootstrap-Based Inference for Clustered Errors," *Canadian Journal of Economics/Revue canadienne d'économique*, 56, 839–858.

Wells, S., C. Ravilious, and E. Corcoran (2006): *In the Front Line : Shoreline Protection and Other Ecosystem Services from Mangroves and Coral Reefs*, UNEP-WCMC.

# G Experimental Instructions

## Experimental Instructions

### 1. Introduction

Welcome. My name is ____. This game consists of **a problem-solving game** played in groups of two, in which you and the other player can earn additional money. At the start of the game, you will be given **15 Kina** with which you will play the game. There is a possibility that you will lose this money during the game, but we will explain how you can prevent this from happening.

### 2. Situation

For this game, we kindly ask you to imagine the following situation: You live on an island that is heavily affected by climate change and that is harmed more and more often by extreme weather events. For example, think back at the Cyclone Ita (in April 2014) which was a storm that caused widespread damage in Bougainville. The storm brought heavy rain, coastal flooding, and strong winds that caused significant damage to infrastructure, homes, and crops. Or think about the sea level that continues steadily and that already forced some entire villages to relocate to further inland. Lastly, think about the flooding (in December 2018) in many parts of Bougainville, damaging homes, infrastructure, and crops. The flooding also disrupted transportation and caused landslides, which further compounded the damage.

It is known and has been widely shown that because of climate change, these extreme weather events happen more frequently and more severely. In order to adapt to these threats that become stronger but also more unpredictable over the years, communities can take action. You can coordinate in order to protect yourself better for those situations in the future. This way you can decrease the caused harm as much as possible. Precisely, by coordination, we mean village members can engage in community funds that manage reserve or prevention mechanisms for the sake of upcoming disasters. On a bigger scale, district climate funds could engage in preservation and re-forestation of mangroves in the North-West of Bougainville that would benefit the whole region. Mangroves not only capture carbon emissions and thereby contributing to the slowing down of sea level rise, they also provide a strong under water root system that detains water flooding into the land in times of coastal flooding and tsunamis.

The fund could also be used for sea wall building in those locations where the sea level rise is affecting villages most. Similarly, as for other district or community fund situations that you and others normally engage in, the success of this depends on the cooperation of all. If all engage in the district climate fund, the benefit is bigger and you all benefit from this in the moment this is needed. However, if only some engage in it, the benefit becomes smaller for all whereas the costs are only born by the few people engaging in it.

### 3. Explanation of the Game

For the game that you are about to play you have now been randomly assigned to a group of 2 players (you and someone else). You don't know who the other person is, only that it is a person from the same village as you/from a village in the North-West of Bougainville.

We give you a starting money of 15Kina with which you play the game. With this money, you can protect yourself and further people in your village and Bougainville in general against the threat of a natural disaster.

Please note: This is only an example! The protection holds for the disaster of the game where we hypothetically induce a disaster for which you can protect. This is not an insurance for the real case when a disaster comes in the future.

In the game, you and your respective game partner have two options:

1) *Investing in a district climate fund for protection reasons in case of a disaster happening – Cost: 7Kina,* or
2) *Not investing in a district climate fund. So, there is no protection in case of a disaster happening – No costs*

The card draw decides whether the natural disaster happens and reduces your money or not. So, there is a 50% probability that the disaster occurs and a 50% probability that is does not occur.

If the natural disaster does occur, you lose money. How much you lose depends on the decision taken by you and the other player regarding the climate protection investment. You can be:

- Fully covered: If you have both invested in the district climate fund
- Partially covered: If only one of you has invested in the district climate fund
- Not covered: If none of you has invested in the district climate fund

If the natural disaster does not occur, you do not lose anything. If you have invested into the district climate fund, you will lose this investment.

When you take your decision, you cannot communicate with the other player. You will therefore not know what your game partner decides at the moment when you have to take a decision.

**4. Examples**

For example, if the disaster does not happen, there are two possible outcomes for you:

(1) You have not invested into the district climate fund. You keep all your ***15Kina***
(2) You have invested into the district climate fund which has a cost of 7 Kina. So, you keep ***8 Kina***

However, if the disaster does happen, there are different potential outcomes for you depending on what you and the other player did:

(1) **Fully protected:** You have both invested into the district climate fund. You have invested 7 Kina for the district climate fund but the rest of your money is protected, so you keep ***8 Kina***
***(2)*** **Partially protected through you:** Only you have invested into the district climate fund but your group member did not. You have invested 7 Kina for the district climate fund and you are both partially protected, meaning the disaster destroys 8 Kina of your money. So, you lose all your money: ***15-7-8=0 Kina***
(3) **Partially protected through game partner:** Only your group member has invested into the district climate fund but you did not. You are both partially protected, meaning the disaster destroys 10 Kina of your money. So, you keep ***15-8=7 Kina***
(4) **Not protected:** You have both not invested into the district climate fund. You are not protected at all and lose 15 Kina, so you keep ***15-15=0 Kina***

**5. Situation and Decision**

Imagine that the forecasts predict that there is a 50% chance that a natural disaster will occur on the island. If the tsunami occurs, coastal flooding will enter your and other coastal villages. You know that mangrove trees at the coast lines reduce the wave energies of tsunamis and coastal flooding which can save the crops.

You in your community and the other communities of the island can take action by engaging in some kind of protection. You now have to decide whether or not you want to invest into the district climate fund devoted to this activity. For this activity, you get two envelops: One is for the money that you keep for yourself and one is for the district climate fund in case you decide to pay for it.

If you want to invest into the district climate fund, you put 7Kina into the 'district climate fund' envelop and the rest (8Kina) into the envelop for the money that belongs to you. If you don't want to invest into the district climate fund, you put nothing into the 'district climate fund' envelop and all 15Kina into the envelop for the money that belongs to you.

Remember that if you don't invest into the district climate fund and your game partner also doesn't invest into the district climate fund, you both have no protection at all. That means if the disaster comes next year, the money of you and your game partner will be destroyed, so you are left with 0Kina. If the disaster does not occur, you can both keep all the money, so 15Kina.

Remember that if only one of you two invests into the district climate fund, you or your game partner, you only have partial protection against the disaster. That means if the disaster occurs, the money of you and your game partner will be reduced by 10Kina. So, it will be 5Kina for the person who has not invested into the district climate fund and 0Kina for the person who has invested.

If the disaster does not occur, no money will be lost. The person who has not invested into the district climate fund keeps all the money, so 15Kina, and the person who has invested into the district climate fund with 5Kina keeps 10Kina.

Remember that if both of you invest into the district climate fund, you and your game partner, you have full protection against the disaster. That means no matter if the disaster comes or not next year, the money of you and your game partner will be saved and you keep all the money, so 15Kina minus the payment of the investment (5Kina), so 10Kina.

I would now like to ask you some questions to check whether you have understood the rules of the game. Then, I ask you some final questions, next you take your decision while I leave the room before we will finally draw the card that determines whether the disaster will occur or not. Your game partner from your village/a village from another constituency in the North-West of Bougainville has to make the same decision as you. You do not know who this is and you do not know what he/she decides.

**6. Control Questions**

*1)* How many people are you playing the game with? *(Correct answer: 1)*
O 1 other person from my village/from a village in the North-West of Bougainville.
O Alone
O 2 other people from my village/ from a village in the North-West of Bougainville.
O 3 other people from my village/ from a village in the North-West of Bougainville.

2) How much money do you have at the beginning of the game? *(Correct answer: 15)*
O 10 Kina O 5 Kina O 15 Kina O 20 Kina

3) How much money does it cost you to invest into the district climate fund? *(Correct answer: 7)*
O 10 Kina O 7 Kina O 15 Kina O 20 Kina

4) What is the likelihood of the disaster happening? *(Correct answer: Half of the times)*
O Never O Very few times O Half of the times O Always

5) How much money do you have if you have invested into the district climate fund and the disaster did not occur? *(Correct answer: 8)*
O 8 Kina O 5 Kina O 15 Kina O 20 Kina

6) How much money do you have if you have not invested into the district climate fund and the disaster did not occur? *(Correct answer: 15)*
O 10 Kina O 5 Kina O 15 Kina O 20 Kina

7) How much money do you have if you and your group member have not invested into the district climate fund and the disaster did occur? *(Correct answer: 0)*
O 10 Kina O 5 Kina O 15 Kina O 0 Kina

8) How much money do you have if you have not invested into the district climate fund but your group member did and the disaster did occur? *(Correct answer: 7)*
O 10 Kina O 7 Kina O 15 Kina O 2 Kina

**9) Decision**

I everything alright? Do you have any further questions?

If not, I will leave the table while you take your decision. Afterwards I will come back and I will draw a card that determines whether or not the climate disaster will come next year.

Are you done? If yes, I will now return into the room.

**10) Belief elicitation**

I would now like to ask you some questions before conducting the disaster card draw.

1. Do you think the card will draw that the disaster occurs?
   O Yes O No

2. Think about the game that you played with someone else. Do you think she/he has contributed to the district climate fund?
   O Yes O No

3. Do you think the other person expects you to contribute?
   O Yes O No

4. Suppose the big man was playing this game with someone from his village. Do you think he would contribute to the district climate fund?
   O Yes O No

5. Suppose the big man was playing this game with someone from a village from another constituency. Do you think he would contribute to the district climate fund?
   O Yes O No

6. Suppose the big woman was playing this game with someone from her village. Do you think she would contribute to the district climate fund?
   O Yes O No

7. Suppose the big woman was playing this game with someone from a village from another constituency. Do you think she would contribute to the district climate fund?
   O Yes O No

8. Suppose this envelop is from another man in your village who plays the game with you. Do you think he contributes to the district climate fund?
   O Yes O No

9. Suppose this envelop is from another man from a village of another constituency who plays the game with you. Do you think he contributes to the district climate fund?
   O Yes O No

10. Suppose this envelop is from another woman in your village who plays the game with you. Do you think he contributes to the district climate fund?
    O Yes O No

11. Suppose this envelop is from another woman from a village of another constituency who plays the game with you. Do you think she contributes to the district climate fund?
    O Yes O No

12. Suppose this envelop is from another man from a village of another constituency who plays the game with someone from his village. Do you think he contributes to the district climate fund?
    O Yes O No

13. Suppose this envelop is from another man from a village of another constituency who plays the game with someone from your village. Do you think he contributes to the district climate fund?
    O Yes O No

14. Suppose this envelop is from another woman from a village of another constituency who plays the game with someone from her village. Do you think she contributes to the district climate fund?
    O Yes O No

15. Suppose this envelop is from another woman from a village of another constituency who plays the game with someone from your village. Do you think she contributes to the district climate fund?
    O Yes O No

16. What do you think most of the people in your village do?
    O Contribute to the district climate fund
    O Don't contribute to the district climate fund

17. What do you think most of the people from villages in other constituencies do?
    O Contribute to the district climate fund
    O Don't contribute to the district climate fund

18. What do you think does the other player in your game?
    O Contribute to the district climate fund
    O Don't contribute to the district climate fund

19. What do you think your village wants you to do?
    O Contribute to the district climate fund
    O Don't contribute to the district climate fund

20. What do you think the big man/big woman wants you to do?
    O Contribute to the district climate fund
    O Don't contribute to the district climate fund

21. Are you afraid of doing something against the will of the big man/big woman?
    O Yes
    O No

22. Do you believe a district climate fund would be beneficial for your village?
    O Yes
    O No

23. Do you believe a community climate fund would be beneficial for your village?
    O Yes
    O No

**11) Disaster Decision Outcome and End of the Experimental Session**

Thank you for your answers. You can now draw a card that determines whether the disaster will come or not.

*Card draw that determines the disaster happening or not.*

The card draw has decided that the climate disaster will (not) occur. Your final payout also depends on the decision of the other player. So, you will receive your money at the end of the whole experiment, so once all players in all villages have taken their decision. So, today you take your participation fix fee of 2Kina once you have finished the questionnaire and in two weeks we will make the final payout of the game.

You can now leave this first part of the game and go to the final questionnaire.

**Questionnaire**

**1. Demographic Questions**

1. Please indicate you participant ID: _____

2. Please confirm your gender?
   O Man O Female

3. How old are you?: _____ *years*

4. How many completed years of education do you have?: _____ years

5. Normally how much money do you earn within each week?: _____ kina

6. How often do you go to church?
   O Every day O Several times a week O Once a week O Once a month
   O Once a year O Never

7. What is your main activity in terms of work?
   O Self-sustained O Own plantation O Selling food on the market O Employed
   O Student O Teacher O Other: _____

8. Do you live here since you were born or after that?
   O Yes, I have lived here since I was born O No, I came here after I was born.

   8b. If no, for how many years have you been living here?: _____ years

9. How much money have you spent to recharge your phone credit over the last month?
   O K0 O K5 O K10 O K30 O K100 O More than K100

10. How often have you accessed the internet, Facebook, whatsapp, over the last month?
    O Every day O Several times a week O Once a week O Once a month O Never

11. What is your mother tongue?
    O Location language O Tok Pidgin O Other: ____

12. What language group are you in?: (Halia, Saposa, Selau): _____

13. Do you engage (with money, time or advice) in community activities/ community funds?
    O Yes O No

    13c. If yes, how often?
    O Every day O Several times a week O Once a week O Once a month
    O Once every six months O Once a year

## 2. Social preference questions and reputational concerns

1. Do you think people are mostly looking out for themselves as opposed to trying to help each other?

People mostly looking out for themselves

People trying to help each other

1 2 3 4 5 6 7 8 9 10

2. Do you think people would try to take advantage of them if they got a chance as opposed to trying to be fair?

People take advantage of other people

People Try to be fair

1 2 3 4 5 6 7 8 9 10

3. Do you think most people can be trusted or that one needs to be very careful when dealing with people?

O Yes
O No

4. How much do you trust people in general?

O I trust fully
O I rather trust
O I rather distrust
O I don't trust at all

5. How much do you trust people you just met?

O I trust fully
O I rather trust
O I rather distrust
O I don't trust at all

6. Do you generally see yourself as fully prepared to take risks as opposed to generally trying to avoid taking risks?

I am fully prepared to take risks

I rather try to avoid taking risks

1 2 3 4 5 6 7 8 9 10

7. Please think of the game you just took part in. Suppose a person like the Big Man was observing you while you were making your decision. How do you think you would get a better name for yourself?
O Not contributing to the district climate fund gives me a good name
O Contributing to the district climate fund gives me a good name
O My decision will never change my name.

8.Which action would be better for you?
O I contribute to the district climate fund.
O I keep all the money.

9. Which action would be better for the community?
O I contribute to the district climate fund.
O I keep all the money.

**3. Climate Change Questions**

1. Many people nowadays talk about climate change. When you hear the word climate change, what do you think of? _____

2. Do you believe in climate change? For example, do you believe in that the temperatures are rising, that more severe natural disasters happen with more frequency?
   O Yes
   O No

3. Do you think there should be climate change community funds that protect the whole village when climate disasters appear?
   O Yes
   O No

4. Have you ever heard of any of the following natural hazards or environmental changes?
   1. Flooding/Intense rainfall:
      O Yes
      O No
   2. Tsunamis:
      O Yes
      O No
   3. Earthquakes:
      O Yes
      O No
   4. Sea level rise:
      O Yes
      O No
   5. Intense drought:
      O Yes
      O No

6. Mountain erosion (unrelated to heavy rainfall):
   O Yes
   O No
7. Change in weather patterns:
   O Yes
   O No
8. Food scarcity/Failed harvest (due to flood, drought, etc.):
   O Yes
   O No
9. Other: _____

5. How important are the following issues for Bougainville, in your opinion? (*Please indicate one of the options for each: Important, Somewhat important, Unimportant, I Don't know)*
   1. Flooding/Intense Rainfall: _____
   2. Tsunamis: _____
   3. Earthquakes: _____
   4. Sea level rise: _____
   5. Intense drought: _____
   6. Mountain erosion (unrelated to heavy rainfall) : _____
   7. Change in weather patterns: _____
   8. Food scarcity/Failed harvest (due to flood, drought, etc.): _____
   9. Other: _____

6. Has any of the following natural hazards or environmental changes ever harmed you or your family?
   1. Flooding/Intense rainfall:
      O Yes
      O No
   2. Tsunamis:
      O Yes
      O No
   3. Earthquakes:
      O Yes
      O No
   4. Sea level rise:
      O Yes
      O No
   5. Intense drought:
      O Yes
      O No
   6. Mountain erosion (unrelated to heavy rainfall):
      O Yes
      O No
   7. Food scarcity/Failed harvest (due to flood, drought, etc.):
      O Yes
      O No
   8. Other: _____

7. Which of the natural hazards is the most dangerous for your village? (*Please name one only)*
    O Tsunami
    O Earthquakes
    O Intense drought
    O Intense rainfall
    O Mountain erosion (unrelated to heavy rainfall)
    O Change in weather patterns
    O Food scarcity/Failed harvest (due to flood, drought, etc.)

8. Which of the natural hazards is the second most dangerous for your village? (*Please name one only)*
    O Tsunami
    O Earthquakes
    O Intense drought
    O Intense rainfall
    O Mountain erosion (unrelated to heavy rainfall)
    O Change in weather patterns
    O Food scarcity/Failed harvest (due to flood, drought, etc.)

9. Do you feel safe to live in your village with regard to environmental hazards, named above?
    O Yes
    O No

    9b. If no, why? _____

10. Do you act in any way to protect yourself or your belongings from any of the environmental hazards named above?
    O Yes
    O No

    10b. If yes, how? Through:
    O Flood prevention measures (individual level)
    O Flood prevention measures (individual level)
    O Drought prevention measures (individual level)
    O Drought prevention measures
    (community level)
    O Earthquake/Tsunami detection/ prevention measures (community level)
    O Preserving the environment
    O Engaging in local politics matters
    O Other _____

    10c. If no, why not?
    O No time
    O No money
    O No interest
    O Not important enough
    O Other:_____

11. Have you ever feared any natural hazards or environmental changes for food security reasons? *(E.g. droughts or floods)*
    O Yes
    O No

    11b. If yes, which natural hazard causing what?: _____
_____

12. Suppose the government could invest a fixed amount of money in one of the following areas, in order to improve the danger of the natural hazards. Which one would you choose for your community? (*Please indicate only one option)*
    1. Re-forestation to attenuate the floods
    2. Water tanks for reservation for drought periods
    3. Walls for times of beach flooding and generally sea level rise
    4. Other: _____

13. Do you trust the following politicians to implement the correct measures to develop the island and/ or community? (*Please rate each by: Mostly yes, Mostly no, Don't know)*
    1. National Politicians: _____
    2. Regional Politicians (ABG Goverment/ HOR): _____
    3. COC/ COE/ Ward Members: _____
    4. Local Politicians (Big Men/ Big Women): _____
       _____

**4. Vignettes**

1. Story: Fred plants trees

In a distant community that is very much like this community, Mr Fred receives money from a non-governmental organization to plant 20 trees. The trees that are planted in the mountains close to the village help to protect villages against flooding.

Mr Fred has four options:

A) Plant all 20 trees in his own village
B) Plant 10 trees in his own village and leave the other 10 for the two closest villages
C) Plant 4 trees in his own village and leave the other 16 for other villages in the northern area of Bougainville.
D) Plant 1 trees in his own village and leave the other 19 for other villages over the whole of Bougainville.

All other places would also need trees against climate change.

What do you think Mr Fred should do?: A, B, C, D?

2. Story: Thelma and Iris

In a distant community that is very much like this community, there are two women, Thelma and Iris. Thelma and Iris start working on two gardens of the same size and they work about the same time and put in the same effort. After a month of work, a big storm comes and destroys nearly all the harvest of Iris. Thelma's harvest survives entirely and she sells it on the market and she earns a lot of money. Iris has very little to sell.

A) Do you think it is right that Thelma earns much more money than Iris at the end of the month?
   O Very fair
   O Rather fair
   O Rather not fair
   O Not fair at all

B) Thelma decides not to share any harvest or money with Iris. Another person, Richard, gossips about what Thelma did. Do you think that Richard will get a good name, a bad name, or nothing at all?
   O good reputation
   O bad reputation
   O no reputation change

C) Another person, Adrian punishes Thelma. Do you think that Adrian will get a good name, a bad name, or nothing at all?
   O good reputation
   O bad reputation
   O no reputation change

3. Story: Steven and Peter

In a distant community that is very much like this community that has introduced a climate community fund, there are two men, Steven and Peter. Steve has contributed with time and money into the community fund, Peter has not contributed anything. A natural disaster happens and thanks to the climate community fund the harvests of both, Steven and Peter, are saved.

A) Do you think it is right that Peter has not contributed to the climate community fund?
   O Very fair
   O Rather fair
   O Rather not fair
   O Not fair at all

B) The community thinks about excluding Peter from the benefits of the climate change community fund if Peter continues to deny a contribution. Do you think it is right of the community to exclude Peter from the benefits in case a natural disaster happens?
   O Very fair
   O Rather fair
   O Rather not fair
   O Not fair at all