DISCUSSION PAPER SERIES

53896-1749035389

GROWING COOPERATION

Georg Kirchsteiger, Tom Lenaerts and Remi Suchon

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

GROWING COOPERATION

Georg Kirchsteiger, Tom Lenaerts and Remi Suchon

Discussion Paper 53896-1749035389
Published N/A
Submitted 04 June 2025

Centre for Economic Policy Research
187 boulevard Saint-Germain, 75007 Paris, France
2 Coldbath Square, London EC1R 5HL
Tel: +44 (0)20 7183 8801
www.cepr.org

GROWING COOPERATION

Abstract

Experimental evidence shows that in a repeated dilemma setting cooperation is more likely to become the norm in small matching groups than in large ones. This result holds even if cooperation is an equilibrium outcome for all investigated group sizes. But what happens if small matching groups are merged to become large ones? Our paper is based on the idea that due to norm spillovers, a large group created by a merger of small groups is more likely to cooperate than a large group of similar size that is created directly. We tested this idea experimentally in the context of an infinitely repeated prisoner's dilemma game. We compared the cooperation behavior of groups that result from mergers of smaller groups with the cooperation behavior of groups with constant group size. We found that cooperation levels were significantly higher in large groups that resulted from gradual growth than in large groups of the same size that were directly created. Looking at the individual behavior, we see that more subjects develop a norm of unconditional cooperation when the group size increases than when it is already large from the beginning. Hence, our results confirm the idea that cooperation is much more likely to be achieved when groups grow from small to large than when large groups are formed directly.

JEL Classification: C73, C92, D23, D90, L22

Keywords: Prisoner's dilemma, Cooperation in repeated games, Group growth, Norm spillover

Georg Kirchsteiger - georg.kirchsteiger@ulb.be
ECARES, Université Libre de Bruxelles, Bruxelles, Belgium and CEPR

Tom Lenaerts - tom.lenaerts@ulb.be
Machine Learning Group, Université Libre de Bruxelles & Artificial Intelligence Laboratory, Vrije Universiteit Brussel

Remi Suchon - remi.suchon@univ-catholille.fr
ANTHROPO LAB – ETHICS EA 7446, Université Catholique de Lille

Acknowledgements
We thank Antoine Deplancke and Quentin Victeur for helping with the experiment. For useful comments, we thank Marie Claire Villeval, and audiences at GATE-Lab in Lyon, at ASFEE in Grenoble, and at the Universities of Gothenburg and Hamburg. This project received financial support from the Fonds de la Recherche Scientifique (FNRS, PDR 31257234) and the Région Hauts-de-France

Growing cooperation.*

Georg Kirchsteiger†, Tom Lenaerts‡, Rémi Suchon§

June 4, 2025

Abstract

Experimental evidence shows that in a repeated dilemma setting cooperation is more likely to become the norm in small matching groups than in large ones. This result holds even if cooperation is an equilibrium outcome for all investigated group sizes. But what happens if small matching groups are merged to become large ones? Our paper is based on the idea that due to norm spillovers, a large group created by a merger of small groups is more likely to cooperate than a large group of similar size that is created directly. We tested this idea experimentally in the context of an infinitely repeated prisoner's dilemma game. We compared the cooperation behavior of groups that result from mergers of smaller groups with the cooperation behavior of groups with constant group size. We found that cooperation levels were significantly higher in large groups that resulted from gradual growth than in large groups of the same size that were directly created. Looking at the individual behavior, we see that more subjects develop a norm of unconditional cooperation when the group size increases than when it is already large from the beginning. Hence, our results confirm the idea that cooperation is much more likely to be achieved when groups grow from small to large than when large groups are formed directly.

JEL codes: C73, C92, D23, D90, L22.

Keywords: Prisoner's dilemma, Cooperation in repeated games, Group growth, Norm spillover.

---

*We thank Antoine Deplancke and Quentin Victeur for helping with the experiment. For useful comments, we thank Marie Claire Villeval, and audiences at GATE-Lab in Lyon, at ASFEE in Grenoble, and at the Universities of Gothenburg and Hamburg. This project received financial support from the Fonds de la Recherche Scientifique (FNRS, PDR 31257234) and the Région Hauts-de-France

†ECARES, Université Libre de Bruxelles, Bruxelles, Belgium, and CEPR - georg.kirchsteiger@ulb.be

‡Machine Learning Group, Université Libre de Bruxelles & Artificial Intelligence Laboratory, Vrije Universiteit Brussel, Brussels, Belgium - tom.lenaerts@ulb.be

§ANTHROPO LAB – ETHICS EA 7446, Université Catholique de Lille, F-59000 Lille, France. - remi.suchon@univ-catholille.fr

1 Introduction

Humans are social animals that interact in groups. In many situations they can achieve better outcomes for all group members when they cooperate than when they refrain from cooperation. Norms of cooperation, defined as a shared understanding that cooperation is the right thing to do (Bicchieri, 2005; Fehr and Schurtenberger, 2018), are crucial for the success of individuals as well as firms or other institutions. This is particularly true when the strategic situation allows for cooperation as well as non-cooperation as equilibrium outcomes. In such situations some groups develop norms or habits that lead to cooperative outcomes, while others get stuck in non-cooperation. This has been shown in the context of repeated prisoners' dilemma games (for an overview of the experimental results, see Dal Bó and Fréchette (2018)). A robust finding is that such norms of cooperation hardly emerge in large groups with random matching. This is problematic since many important issues require developing large-scale cooperative norms, e.g. the organization of labor in large firms.

In this paper, we study whether gradual growth helps groups to develop cooperative norms that would allow for cooperation in the resulting large groups. In the context of the infinitely repeated prisoner's dilemma, we experimentally test whether cooperation is more frequent in large matching groups when the groups have started small before merging into larger ones. The mechanism that we have in mind is as follows. Small groups are more favorable to the emergence of cooperative norms than large ones: in the prisoner's dilemma, partner matching typically elicits more cooperation than stranger matching, and in stranger matching, larger matching groups elicit less cooperation (e.g. Ghidoni et al., 2019). This is also true in situations when an equilibrium supporting cooperation exists for both group sizes (see Duffy and Ochs, 2009). Because of norm spillovers, a large matching group formed by merging small cooperative groups should be more likely to sustain cooperation than a group that starts out large. In other words: In order to establish cooperation, agents should first interact in small groups to learn a cooperation norm. This cooperation norm will continue to influence the agents' behavior after the small groups get merged into a larger one. This mechanism implies more cooperation in large groups formed by the merger of small groups than in groups that started out large from the beginning.

The emergence of cooperative norms in social dilemmas has been extensively studied (Fehr and

Fischbacher, 2004; Dal Bó and Fréchette, 2018). But the question of whether managing group growth favors the emergence of a norm of cooperation remains relatively unexplored. This is surprising, since starting small is an intuitive way to favor cooperative norms that would spill over to larger groups. For instance, in the workplace teams may start small to build cooperative norms before growing larger. In contrast to the existing literature, the main focus of our paper is not the impact of group size on cooperation. Rather, we investigate how managing the growth of groups *itself* can help build a norm of cooperation. Some evidence suggests that managing growth helps to sustain mutually beneficial outcomes in groups in coordination games (Weber, 2006). However, situations requiring coordination are very different from dilemma situations where cooperation is particularly beneficial as well as particularly hard to sustain. To the best of our knowledge, no previous study has investigated how gradual growth of the group favors the emergence of cooperative norms sustaining cooperation in resulting large groups in the prisoner's dilemma context.

To study the impact of gradual growth experimentally, we used a repeated prisoner's dilemma game with a random stopping rule with an exogenously given stopping probability. The experiment consisted of four treatments, one with growing group sizes (treatment G), and three baseline treatments with constant group sizes (treatments $B2$, $B4$, and $B8$). The G treatment consisted of three parts with different group sizes. In part 1, each subject played an infinitely repeated prisoner's dilemma game with the same partner. At the end of each round, a random mechanism decided whether the pair received a "stop signal". After having received three stop signals, the repeated game ended and the pair moved on to part 2. In part 2, two randomly chosen part-1-pairs merged to form a group of four. In each round of part 2, each participant was randomly and anonymously matched with another member of her/his part-2-group to play the prisoner's dilemma game. After each round the random mechanism decided whether the group received a stop signal. Similar to part 1, part 2 ended when the group received three stop signals. In part 3, two randomly chosen part-2-groups merged to form a group of eight. In each round of part 3, each participant was randomly and anonymously matched with another member of her/his part-3-group to play the prisoner's dilemma game. Again, after each round the random mechanism decided whether the group received a stop signal, and after receiving three stop signals the experiment ended for the group.

In the $B8$ treatment, a session started with the random formation of groups of eight participants. Again, the experiment consisted of three parts. In each round of each part, each participant was randomly and anonymously matched with another member of her/his group to play the prisoner's dilemma game. After each round, the random mechanism decided whether the group received the stop signal. After receiving three stop signals the group moved to the next part of the experiment (except at the end of part 3, when the experiment ended) without any change in the group composition[1]. Therefore, each of the three parts of the $B8$ treatment was equivalent to part 3 of the G treatment. The $B4$ treatment was similar to $B8$, except that groups of four participants were randomly formed at the beginning of the session. Therefore, each of the three parts of the $B4$ treatment was equivalent to part 2 of the G treatment. In the $B2$ treatment, a session started with the random formation of pairs. Each pair played the prisoner's dilemma game in partner matching across the three parts, organized in exactly the same way as in $B8$ and $B4$. Therefore, each of the three parts of the $B2$ treatment was equivalent to part 1 of the G treatment.

In our experiment, information was limited in two ways. First, histories were private information: participants observed only the outcome of the interactions they were involved in. Second, interactions were anonymous: participants did not know the identity of their counterpart and could not track them across different encounters (except of course for the $B2$ treatment and part 1 of the G treatment). This excludes direct reciprocity as soon as the group size was at least 4. We chose to impose such information restrictions to have a setup where cooperation is particularly difficult to achieve, to stress-test our mechanism (Camera et al., 2012; Dal Bó and Fréchette, 2018).

Norm spillovers are particularly important in contexts where cooperation as well as non-cooperation are equilibrium outcomes. We chose the parameters (payoffs and stopping probability) such that full cooperation, full non-cooperation, and anything in between were equilibrium outcomes for all possible group sizes. Hence, the actual outcome depended crucially on the cooperation norm or habit that was established by the respective group. The drawback of such a parameter choice is that standard game theory does not provide any prediction concerning

[1]Obviously, this is equivalent to not dividing the experiment into three parts, but instead requiring 9 stop signals for the whole experiment to end. But we prefer this presentation with three parts in order to simplify the comparison with the G treatment where the three parts differed in group size.

the expected cooperation levels. To overcome this problem, we developed a very simple model that is based on the assumption that cooperation is more likely to become a norm when the group is small, and that there is norm spillover between the three parts of the experiment. This allowed us to derive predictions for the the differences in the cooperation levels of the different treatments caused by norm spillovers.

In accordance with the predictions of our model, in all three parts of the experiment cooperation rates were larger in the $B2$ than in all other treatments. In all three parts cooperation rates were smaller in the $B8$ than in all other treatments. Most importantly, we found significantly more cooperation and cooperative outcomes in part 3 of the G treatment than in the last part of the $B8$ treatment, although in both cases the group size was the same. Looking at the individual behavior we found that this difference was driven by more participants developing a norm of unconditional cooperation when the group size increased to eight than when it was eight already from the beginning. These results confirm our idea that growing groups are more likely to establish cooperation norms than groups that start out large.

The rest of the paper is organized as follows: Section 2 reviews the relevant literature. Section 3 introduces our design. Section 4 presents a simple model and the predictions derived from it. Section 5 presents our results, and Section 6 concludes.

2 Related literature

We contribute to three streams of the literature.

Experiments on cooperation in infinitely repeated games. In theory cooperation as well as non-cooperation can be an equilibrium behavior in the repeated prisoner's dilemma game if the relative payoff of mutual cooperation is high and/or the discount factor is large enough. This result holds not only for partner matching, i.e. for fixed pairs of players (Friedman, 1971). It also holds for stranger matching, i.e. when in every round every member of a fixed matching group gets randomly re-matched with (possibly another) member of his group (Kandori, 1992; Ellison, 1994). Experiments on cooperation in infinitely-repeated prisoner's dilemma games are surveyed in Dal Bó and Fréchette (2018). An important result is that, in contexts where cooperation as well as defection are equilibrium behavior, cooperation was often observed in partner matching but seldom in stranger matching. This was especially true in "information-starved" contexts like ours (Duffy and Ochs, 2009; Honhon and Hyndman, 2020), where participants observed only the outcome of the interactions they were involved in and could not track the behavior of their counterpart across encounters. We contribute to this literature by proposing growth of the matching group as a way to allow large groups to cooperate more.

Group size and cooperation. This literature asks whether smaller or larger groups are more favorable to cooperation. Some experiments found a negative effect of the size of the matching group on cooperation in stranger matching prisoner's dilemma both with infinite (Duffy and Ochs, 2009) and finite repetitions (Ghidoni et al., 2019). Similar results emerged in n-player prisoner's dilemma Cournot games (Huck et al., 2004) as well as in the helping game (Camera et al., 2013; Bigoni et al., 2019). In the context of the linear public good game, several papers reported a weak but positive effect of group size on cooperation (Isaac and Walker, 1988; Nosenzo et al., 2015).[2] However, the structure of the public good games tested in this literature was very different from our setup: First, in these papers the time horizon was finite and therefore cooperation was never an equilibrium behavior, whereas in our setup cooperation was an equilibrium behavior (as well as non-cooperation). Second, in these public good studies all members of a group interacted in every round, allowing for reputation building. In our setting the information structure was such that reputation building was not possible as soon as the matching group was of size 4. Third, in public good games changing the group size impacts either the social surplus generated by a given individual contribution level, or it changes the individual contribution incentives, or both. In our context, the surplus generated by an individual cooperation decision did not depend on group size and therefore the individual incentives to cooperate were not affected by group size. This allowed us to cleanly focus on the impact of the group size on the emergence of cooperative norms without confounding effects of reputation building and changes in the payoff structure. We contribute to this literature by confirming the negative effect of the matching group on cooperation. More importantly we explore the effect of the dynamics of group size and show that the effect of group size can be

[2] Nosenzo et al. (2015) nuanced this by showing that in setups conducive to cooperation, the group size effect was negative.

mitigated when groups grow instead of starting out large.

A few papers studied the effect of a gradual growth of the scale of interactions on social outcomes. Weber (2006) studied a coordination game and showed that allowing groups to grow slowly allowed them to solve miscoordination issues often observed in large groups. In contrast, our focus is on cooperation. In our game, both cooperation and defection could be sustained as equilibrium outcomes of the repeated game, and mutual defection was the only equilibrium of the stage game. This made our setting more challenging. A few experiments tested the effect of growing groups using finitely repeated linear public good games (Charness and Yang, 2014; Ranehill et al., 2014). Ranehill et al. (2014) found that slow group growth sustained more cooperation than fast growth, which is consistent with our results. However, our study is quite different from theirs: In their game, the social payoff from cooperation increased as the group grew, which makes it hard to pinpoint the pure effect of the growth rate. Furthermore, in our game infinite repetition allowed for cooperation to be played in equilibrium, which was not the case in their game. Despite that theoretical result, cooperation was more commonly observed in such public good games than in infinitely repeated stranger matching prisoner's dilemma. Therefore our results were found in a setup tougher for cooperation. Finally, recall that the effect of group size on public good contribution was more nuanced, and in particular tended to be positive when contributions generated more collective surplus as the group grew.

The emergence and spillover of norms of cooperation Third, we contribute to the literature on the emergence of cooperative norms in social dilemmas (see e.g Ostrom, 2000; Bicchieri, 2005, for reviews). The bulk of the literature focuses on institutions or mechanisms such as punishment, reward, communication, reputation or assortative matching (see e.g Fehr and Fischbacher, 2004; Dal Bó and Fréchette, 2018; Fehr and Schurtenberger, 2018, for reviews). We contribute to the literature on norm spillovers in situations in which a norm of cooperation developed in a context conducive to cooperation spills over to other, possibly less conducive contexts. It was shown that cooperative norms that developed outside of the lab translated into more cooperation in the lab. In Gneezy et al. (2016) participants, who experienced a more cooperative workplace in their everyday life developed more cooperative norms and tended to cooperate more in the lab in one-shot trust games. Molina et al. (2023) investigated the impact of local norm enforcement on the likelihood of cooperating with complete strangers in a one-shot PD. Closer to our paper, some experiments study whether participants who were experimentally put in a context favorable to cooperation internalized a norm of cooperation, with positive spillovers on cooperation in a different context. Engl et al. (2021) studied institutional spillovers in the context of a public good game. In their design, participants participated in two public good games simultaneously. In some treatments, punishment was introduced in one of the games, leading to increased contributions in both games. This result indicated the existence of institutional spillovers. Galbiati et al. (2018) investigated the impact of fines as a sanctioning mechanism for non-cooperation and found that these fines had a lasting positive effect on cooperation that persisted after the fines were lifted. Both these results are consistent with the mechanism of norm spillover that we investigate in our paper. In Peysakhovich and Rand (2016) participants played a repeated PD. In some conditions, cooperation was an equilibrium behavior, while in some others, it was not. Participants who were confronted with PD parameters conducive to cooperation cooperated also more in the games where cooperation was not an equilibrium behavior. This indicates that being put in a context conducive to cooperation allows for the emergence of cooperation norms that carry over to subsequent situations not conducive to cooperation. In contrast to these papers, our paper investigates norm-spillovers from small to large groups in situations where cooperation is always one of many possible equilibrium outcomes.

3 Experimental design

To test for the impact of the growth of the group size on cooperation, we ran four different treatments. In the growth treatment G the group size increased, while in the three baseline treatments–$B2$, $B4$, and $B8$–the group size did not change during the experimental session. Subjects were informed that all their interactions in the experiment would be anonymous. In each experimental session, we ran only one treatment, and each subject could participate only in one session. Each experimental session was divided into three parts. We first present the growth and then the baseline treatments.

3.1 The Growth treatment G

At the beginning of part 1 of each session of the G treatment, each subject was randomly matched with another subject. These pairs remained the same during part 1, i.e. each subject interacted with the same partner during the whole part 1 of the experiment. In each round of part 1 each pair played the following prisoner's dilemma game:

	C	D
C	20,20	0,30
D	30,0	10,10

Table 1: Matrix of the Prisoner's' dilemma game.

There was a 10% chance that the subjects received a "stop signal" at the end of every round. After receiving three stop signals, part 1 of the experiment ended, and part 2 (see next paragraph) started. Stop signals were drawn at the level of the session, so all subjects in a session played for the same number of rounds. We decided to require three stop signals before the first part of the experiment ends in order to reduce the differences in the interaction lengths between different sessions.[3]

At the beginning of part 2 of the experiment, each pair was matched randomly with another pair to form a group of four subjects. In each round of part 2, each subject was randomly and anonymously matched with another member of her/his part-2 group to play the same prisoner's dilemma game as in part 1. Again, there was a 10% chance that the subjects of a session received a stop signal at the end of each round. As in part 1, after three stop signals, part 2 ended.

At the beginning of part 3 of the experiment, each part-2 group was matched randomly with another part-2 group to form a group consisting of eight subjects. In each round of part 3, each subject was randomly and anonymously matched with another member of her/his part-3 group to play the same prisoner's dilemma game. Again, there was a 10% chance that subjects received a stop signal at the end of every round. After three stop signals, the session ended.

[3] In the instructions we stated that each part of the experiment consists of three sequences, and each sequence stops after receiving one stop signal. Obviously, this is equivalent to requiring three stop signals before the whole part ends. But we think the description using sequences was easier to understand for the subjects.

3.2 The baseline treatments $B2$, $B4$, and $B8$

In the baseline treatments, the group size was fixed at 2, 4, or 8 subjects throughout the experiment.

At the beginning of each session of the baseline treatment $B2$, each subject was randomly matched with another subject from the session. Each pair stayed together during the whole experiment. In each round each pair played the prisoner's dilemma game. At the end of each round, there was a 10% chance that participants in the session received a stop signal. After receiving three stop signals part 1 of the experiment ended and part 2 started, without any change in the composition of the pairs.[4] Like the G treatment, the $B2$ treatment consisted of three parts.

At the beginning of each session of the baseline treatment $B4$, participants were randomly divided into groups of four participants. Each group stayed together during the whole experiment. In each round, each subject was randomly matched with another member of her/his group to play the prisoner's dilemma game. At the end of each round, participants had a 10% chance of receiving a stop signal. After receiving three stop signals, part 1 of the experiment ended and part 2 started, without any change in the composition of the group.[5] Like the other treatments, the $B4$ treatment consisted of three parts.

Treatment $B8$ was the same as treatment $B4$, except that in $B8$, participants were divided into groups of eight participants. Each group stayed together during the whole experiment.

3.3 Parameters and information structure

To study the effect of gradual growth of the group on cooperation, we chose a setup in which full cooperation is an equilibrium outcome even for groups of 8 (Duffy and Ochs (2009)). But at the same time the folk theorem shows that any other (non)cooperation level is also an equilibrium outcome. In this kind of setup with many possible equilibrium outcomes the development of a norm of cooperation is particularly important. In addition, empirical studies have shown that the chosen parameters elicited low levels of cooperation (Duffy and Ochs (2009)). Therefore, this parameter structure allowed us to stress-test our mechanism in a situation unfavorable

[4]See footnote 1

[5]See footnote 1

to cooperation. Finally, interactions were anonymous and histories were private information in our experiment. This information structure made reciprocation towards specific individuals impossible (except in the first part of the G and in the $B2$ treatment, where partner matching was used). Such information restrictions were shown to be very unfavorable to cooperation in large groups (Duffy and Ochs, 2009; Dal Bó and Fréchette, 2018). Again, we made this choice to stress-test our mechanism in an environment particularly unfavorable to cooperation.

3.4 Procedures

The experiment was developed using z-Tree (Fischbacher, 2007). All sessions were conducted at Anthropo-Lab, Lille, France. We ran 21 sessions in total, 7 G sessions, 3 $B2$ sessions, 4 $B4$ sessions, and 7 $B8$ sessions. 428 participants were recruited. 152 subjects participated in the G, 56 in the $B2$, 76 in the $B4$, and 144 participated in the $B8$ sessions. We had a total of 19 groups of 8 subjects in G, 28 pairs in the $B2$, 19 groups of 4 subjects in the $B4$, and 18 groups in the $B8$ sessions[6]. Table C.1 in Appendix C summarizes the distribution of sessions and participants in the different treatments. The instructions for the whole experiment were handed out to participants at the beginning of each session and read aloud by the same experimenter. The full instructions can be found in Appendix D. Before starting the experiment, participants had to successfully pass an understanding questionnaire. Questions were answered in private.
Sessions lasted between 60 and 90 minutes. All the rounds were paid. The average payoff was €20.2 (min = 11.9, max = 36.7, SD = 4.82), including a €7.5 show-up fee. Payments were made through Lydia, a popular French banking app, shortly after the end of the respective experimental session.

4 Theoretical Predictions

In order to interpret the results of our experiment, we developed a simple model that allows us to show the impact of norm spillovers in the context of our experimental games. Since the folk theorem applies to these experimental games, a standard game-theoretic analysis does not

[6]Note that in the G treatment each part-3 group of 8 subjects provided an independent observation. In the $B2$, $B4$, and the $B8$ treatment each group of 2, 4, and 8 subjects, respectively, generated an independent observation.

provide any predictions. The results of previous experiments suggest that cooperation is more likely to emerge when the group size is small (see, e.g., Duffy and Ochs (2009)). Therefore, our model assumes that more cooperative equilibria are played when the current group size is small than when the current group size is large. Recall that the whole experiment consists of three parts, with different group sizes in each part of the G treatment. We allow for the possibility that the size of future groups has a negative impact on current cooperation rates. The reason is that players are forward-looking. More importantly, if small groups develop more cooperative norms, and if these norms spill over, the size of past groups should have a negative impact on current cooperation rates.

To fix ideas, consider the three functions $c_p : \mathbb{N}^3 \to \mathbb{R}_+$, $p \in \{1, 2, 3\}$. For each part p of the experiment, $c_p(n_1^t, n_2^t, n_3^t)$ denotes the percentage of cooperative choices in treatment t, $t \in \{G, B2, B4, B8\}$, when the group sizes of the three parts of treatment t are n_1^t, n_2^t, and n_3^t. As explained, we assume that c_p is strictly decreasing in all three arguments. This assumption implies:

Proposition

a) In all three parts of the experiment, the cooperation rates are lower in $B8$ than in any other treatment:

$$c_p(n_1^{B8}, n_2^{B8}, n_3^{B8}) < c_p(n_1^t, n_2^t, n_3^t) \quad \text{for all } p \in \{1, 2, 3\}, t \in \{G, B2, B4\}.$$

b) In all three parts of the experiment, the cooperation rates are higher in $B2$ than in any other treatment:

$$c_p(n_1^{B2}, n_2^{B2}, n_3^{B2}) > c_p(n_1^t, n_2^t, n_3^t) \quad \text{for all } p \in \{1, 2, 3\}, t \in \{G, B4, B8\}.$$

Proof a) $n_p^{B8} \geq n_p^t$ for all $p \in \{1, 2, 3\}$ and $t \in \{G, B2, B4\}$. Furthermore, for all $t \in \{G, B2, B4\}$, there exists a $\widehat{p} \in \{1, 2, 3\}$ such that $n_{\widehat{p}}^{B8} > n_{\widehat{p}}^t$. Hence, our assumption about the impact of group sizes guarantees that the cooperation rates are lower in $B8$ than in any other treatment in all three parts of the experiment.

b) $n_p^{B2} \leq n_p^t$ for all $p \in \{1, 2, 3\}$ and $t \in \{G, B4, B8\}$. Furthermore, for all $t \in \{G, B4, B8\}$, there exists a $\widehat{p} \in \{1, 2, 3\}$ such that $n_{\widehat{p}}^{B2} < n_{\widehat{p}}^{t}$. Hence, our assumption about the impact of group sizes guarantees that the cooperation rates are higher in $B2$ than in any other treatment in all three parts of the experiment.

Note that part a) of the Proposition implies that in the third part of the experiment cooperation rates are higher in the G than in $B8$ treatment. An empirical validation of this prediction would support the main idea of the paper that due to norm spillovers a large group created by merging small groups is more likely to cooperate than a directly created large group of the same size.

The model does not provide any prediction about the comparison between the G and $B4$ treatment. This is due to countervailing effects: in part 1, the group size is smaller in the G than in the $B4$ treatment whereas in part 3, the group size is larger in the G than in the $B4$ treatment. These effects are countervailing, in particular in part 2 of the experiment when the current group sizes are the same in both treatments. If the effect of future group size is larger than that of the past group size, part 2 cooperation rates should be larger in the $B4$ than in the G treatment. It should be the other way around if the effect of future group size is smaller than that of past group size. It is also possible that these two effects cancel each other out.

In summary, the model provides three predictions:

Predictions

Prediction 1: Cooperation rates are lower in $B8$ than in any other treatment.

Prediction 2: Cooperation rates are higher in $B2$ than in any other treatment.

Main Prediction: In the third part of the experiment, cooperation rates are higher in G than $B8$.

The next section investigates whether the experimental results support these predictions.

5 Results

5.1 Overview of the results

Table 2 reports cooperation rates and the occurrence of mutual cooperation - where both players cooperate - separated by treatment and part of the experiment. As expected, cooperation

	Part			
Treatment	1	2	3	Total
B8				
Cooperate	0.219	0.158	0.147	0.177
Mutual Cooperation	0.0713	0.0343	0.0317	0.0469
Obs.	5024	4136	4664	13824
B4				
Cooperate	0.344	0.248	0.169	0.255
Mutual Cooperation	0.155	0.0934	0.0526	0.101
Obs.	2144	2656	1976	6776
G				
Cooperate	0.488	0.253	0.249	0.319
Mutual Cooperation	0.374	0.128	0.108	0.193
Obs.	4456	6640	4536	15632
B2				
Cooperate	0.626	0.723	0.755	0.684
Mutual Cooperation	0.528	0.695	0.734	0.625
Obs.	2486	1576	1186	5248
Total				
Cooperate	0.395	0.276	0.246	0.307
Mutual Cooperation	0.26	0.156	0.131	0.184
Obs.	14110	15008	12362	41480

Table 2: Cooperation rates by part, and by treatment.

rates varied substantially between the different treatments. The cooperation rate was 17.7% in the $B8$, 25.5% in the $B4$, 31.9% in the G, and 68.4% in the $B2$ treatment. A similar pattern was observed for the occurrence of mutual cooperation, with 4.69% in the $B8$, 10.1% in the $B4$, 19.3% in the G, and 62.5% in the $B2$ treatment. Overall, the lowest cooperation rate was observed in part 3 of the $B8$ treatment with 14.7%, and the highest rate was observed in part 3 of the $B2$ treatment with 75.5%. Again, the same pattern was observed for mutual cooperation, with 3.17% in part 3 of the $B8$ and 73.4% in part 3 of the $B2$ treatment.

Table 2 also shows that cooperation tended to decline over time in all treatments but the $B2$

treatment, where cooperation actually increased over time. In addition, in the $B4$ and the $B8$ treatments, cooperation decreased both from part 1 to part 2 and from part 2 to part 3. In the G treatment, cooperation dropped when groups expanded from 2 to 4 members, but remained stable as the group size grew further to 8. These results are confirmed by the marginal effects of each part of each treatment separately (see Table A.1 of Appendix A).

5.2 Test of the theoretical predictions

Result 1: Prediction 1 is supported by the data – cooperation rates were lower in $B8$ than in any other treatment.

Support: As already discussed, Table 2 shows that cooperation rates were lowest in the $B8$ treatment. The table also shows that the same holds if we consider each part of the experiment separately. Further support is found in Table 3. This table reports the marginal effects of logit regressions explaining individual cooperation (empirical models (1) to (4)) or the occurrence of mutual cooperation (where both players cooperate, models (5) to (8)) by a categorical variable indicating the treatment. We use standard errors clustered at the highest group level to account for potential intra-group correlation: the group of eight in B8 and G, the group of four in B4, and the pair in B2. Models (1) to (4) use random effects at the individual level. The $B8$ treatment is the reference category. Hence, the table reports the marginal effects of each treatment compared to $B8$. In every model, we control for time trends by introducing a variable controlling for the period, a variable controlling for the number of "stop signals" remaining before the end of the current part (1, 2, or 3), and demographic controls. In models (1) and (5), we use the data from all parts, while the other models restrict the sample to data from one specific part of the experiment. Models (1) and (5) also include a categorical variable indicating the part of the experiment (1, 2 or 3).

The marginal effects are all positive, and mostly significant at the 5 percent level, indicating that all the other experimental conditions elicited higher cooperation levels than the $B8$ treatment. The only exception was part 3 of the $B4$ treatment, for which the marginal effect was positive but not significant. Therefore, the regression analysis provides strong support for Prediction 1.

Result 2: Prediction 2 is supported by the data – cooperation rates were higher in $B2$ than in any other treatment.

Support: Result 2 is also supported by Table 2. It is also confirmed by the regressions: The estimated marginal effects of $B2$ reported in Table 3 are all very large and significant at the 1% level. In addition, the marginal effects of the $B2$ dummies are always significantly larger than those of $B4$ and G, as indicated by the p-values of the Wald tests reported in the last three lines of Table 3. All the p-values are below 0.05 (and most of them below 0.001).
Note that in part 1 of the experiment cooperation levels were significantly lower in G than in $B2$, despite the fact that, in both cases, fixed pairs played the game. This indicates that the anticipated size of the group did indeed matter – subjects were forward looking, as assumed by the theoretical model.

Result 3: The main prediction is supported by the data – in part 3 of the experiment, cooperation was more frequent in G than in $B8$.

Support: Table 2 shows that cooperation was observed in 14.7% of all cases in part 3 of the $B8$ treatment, whereas the cooperation rate was 24.9% in part 3 of the G treatment. In model (4) of Table 3, the estimated marginal effect of G is 9.06% ($p = 0.042$). A similar result is found for the occurrence of mutual cooperation in part 3 of the experiment, observed in 3.17% of the interactions in $B8$ against 10.8% in G. The estimated marginal effect of G is 8.26%, which is again significant ($p = 0.04$).

5.3 Individual behavior

Our main new result is that large groups which result from gradual growth cooperate more. Next, we dig into the individual behavior that underpins this result. We do a simple analysis using the data of part 3 of the experiment. For each participant we compute the proportion of times (s)he cooperated in part 3 after experiencing defection of her/his partner in the previous

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Coop.=1	Coop.=1	Coop.=1	Coop.=1	Mut. Coop.=1	Mut. Coop.=1	Mut. Coop.=1	Mut. Coop.=1
B8	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
B4	0.111***	0.156***	0.0917**	0.0586	0.0603**	0.0896**	0.0553**	0.0327
	(0.039)	(0.048)	(0.045)	(0.037)	(0.024)	(0.045)	(0.028)	(0.027)
G	0.125***	0.242***	0.0885**	0.0906**	0.146***	0.297***	0.0827***	0.0826**
	(0.033)	(0.034)	(0.037)	(0.042)	(0.025)	(0.031)	(0.025)	(0.040)
B2	0.527***	0.420***	0.594***	0.643***	0.556***	0.455***	0.616***	0.726***
	(0.078)	(0.076)	(0.083)	(0.068)	(0.071)	(0.077)	(0.081)	(0.071)
N	41480	14110	15008	12362	20740	7055	7504	6181
Part	All	1	2	3	All	1	2	3
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
p-values of Wald-tests:								
B4 v B2	< 0.001	.002	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
B4 v G	.74	.082	.948	.492	.004	< 0.001	.418	.278
G v B2	< 0.001	.024	0	< 0.001	< 0.001	.047	< 0.001	< 0.001

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors clustered at the highest group level (the group of eight in B8 and G, the group of four in B4, and the pair in B2) are reported in parentheses. Marginal effect from logit models are reported. Models (1) to (4) include random effects at the individual level. Individual controls include age, gender, professional status, monthly expenses, difficulty understanding the instructions and past participation in behavioral experiments. Time controls include include a continuous variable for the current period and a categorical variable indicating how many "stop signals" remain before the end of the current part. In models (1) and (4), we also introduce a variable indicating the part of the experiment.

Table 3: Regressions testing our theoretical predictions.

round. We also compute for each participant the proportion of time (s)he cooperated in part 3 after experiencing cooperation of her/his previous-round partner.[7]

We use a clustering algorithm (k-means) to classify subjects into 3 clusters.[8] Two of the resulting clusters correspond roughly to unresponsive behavior i.e. strategies in which decisions do not depend on the behavior of the last partner: Always cooperate ("AC" cluster), or always defect ("AD" cluster). Overall, 73 subjects are classified in the AC cluster (17.06% of all subjects). These subjects cooperated 96% of the times when the previous partner cooperated and 76% of the times when the previous partner defected. 253 subjects (59.11%) are classified in the AD cluster. These subjects cooperated 2.26% of the times when the previous partner defected and 4.1% of the times when the previous partner cooperated. Note that irrespective of the parameters of the game, AD is always an equilibrium strategy, while AC is never an equilibrium strategy. The remaining cluster is more mixed. 102 subjects (23.83%) belong

[7]This analysis considers only the previous-round decision of the partner as determining a subject's behavior, which is reminiscent of memory-1 strategies. Past literature shows that memory-1 strategies are sufficient to describe most of the observed behavior (Dal Bó and Fréchette, 2018, 2019).

[8]We find similar results when increasing the number of clusters to 4 or 5.

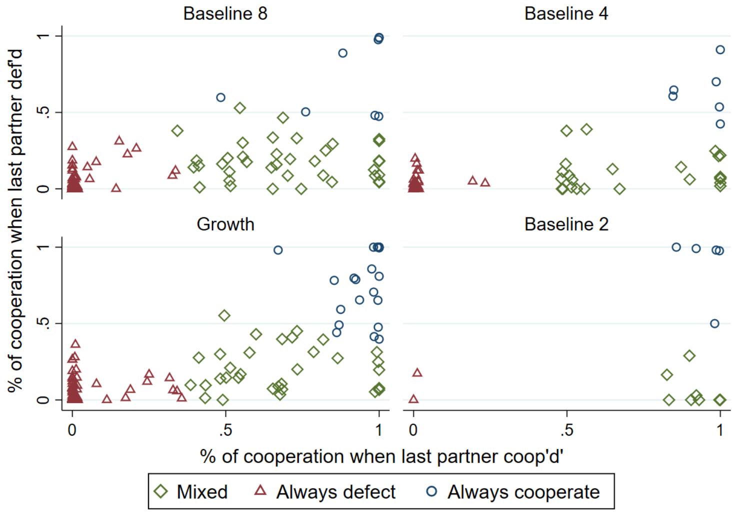

Figure 1: Cluster analysis of individual behavior in part 3 of the experiment.

to this cluster, and these subjects cooperated 71.4% of the times when the previous partner cooperated and 16.1% of the times when the previous partner defected. This indicates that many subjects in the "mixed" cluster were conditional cooperators.

The clustering algorithm was not able to cluster 94 subjects (22.43% of all subjects), because one of the two classification dimensions was missing: In part 3, 32 subjects never experienced defection, and 62 never experienced cooperation. Subjects who never experienced defection cooperated in 99% of all rounds on average (each individual participant in at least 82% of all rounds). Therefore, we classified them in the AC cluster. All of them were in the $B2$ treatment. Subjects who never experienced cooperation cooperated on average 3.9% of all rounds (each of them cooperated in less than 28% of all rounds). They were classified in the AD cluster. These included 22 subjects from the $B8$, 12 from the $B4$, 19 from the G treatment, and 9 from the $B2$ treatment. Obviously, this classification is somewhat arbitrary, and may mistake a conditional cooperator for an unconditional one. But dropping these subjects from the analysis does not change our results.

Figure 1 plots the subjects according to their cooperation rates following defection and their cooperation rates following cooperation, for the four treatments separately. Different markers are used to identify different clusters, blue circles for the AC subjects, red triangles for the AD subjects, and green diamonds for the subjects belonging to the mixed cluster. Figure 1 shows that the distribution of subjects into the clusters was very different for the $B2$ treatment. This is not surprising given the very different nature of interactions with the same partner compared to interactions with changing co-players. Figure 1 also indicates the reason for the different part-3 outcomes of the G and the $B8$ treatment.

Result 4: In part 3 of the experiment, there were significantly more unconditional cooperators in the G than in the $B8$ treatment

Support: Table 4 reports the distribution of the subjects into the three clusters across treatments. 22 subjects from the G treatment (14.47%) are classified as unconditional cooperators, while only 8 AC subjects were observed (5.56%) in the $B8$ treatment. Applying the Fischer exact test reveals that this difference is significant at the 5% level ($p = 0.012$). The percentage of mixed-cluster subjects was the same in the G and the $B8$ treatment. In Appendix B, we show that we get similar results using the Strategy Frequency Estimation Method (used by e.g. Dal Bó and Fréchette, 2011; Fudenberg et al., 2012; Bland, 2020).

	B8	B4	G	B2	Aggregate
AC cluster	5.56% (8)	7.89% (6)	14.47% (22)	66.07% (37)	17.06% (73)
AD cluster	68.05% (98)	60.53% (46)	64.48% (98)	19.64% (11)	59.11% (253)
Mixed cluster	26.39% (34)	31.58% (24)	21.05% (32)	14.29% (8)	23.83% (102)
Total	144	76	152	56	428

Frequencies are in parentheses

Table 4: Distribution of participants on the clusters using the k-mean clustering algorithm.

This result suggests that the positive impact of the growth of the group size on cooperation was driven by subjects who adopted an unconditional cooperation strategy in G. When excluding subjects who belonged to the AC cluster, we find that in part 3 of the experiment cooperation rates were not significantly different in the $B8$ and the G treatment (13.7% in $B8$ and 13.2% in G). This provides further evidence that the AC subjects were the reason for the different part 3 outcomes in G and $B8$.

6 Conclusion

The results of our paper show that cooperation within large groups can be promoted by a gradual growth of the group. When interacting in small groups, agents adopt a cooperation norm. These cooperation norms carry over when smaller groups get merged into larger ones. In our setup, smaller groups were randomly chosen to merge into larger ones. But one can imagine that the choice of which groups merge to form a larger group is not random, but depends on the cooperation levels the small groups achieved. For example, one could use a form of assortative matching where small groups of similar productivity, i.e. similar cooperation levels, merge. Such a merger procedure might increase the impact of organic growth on cooperation. Analyzing the impact of different merger procedures is left to further investigation.

References

Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms.* Cambridge University Press.

Bigoni, M., G. Camera, and M. Casari (2019). Partners or strangers? cooperation, monetary trade, and the choice of scale of interaction. *American Economic Journal: Microeconomics 11*(2), 195–227.

Bland, J. R. (2020). Heterogeneous trembles and model selection in the strategy frequency estimation method. *Journal of the Economic Science Association 6*(2), 113–124.

Camera, G., M. Casari, and M. Bigoni (2012). Cooperative strategies in anonymous economies: An experiment. *Games and Economic Behavior 75*(2), 570–586.

Camera, G., M. Casari, and M. Bigoni (2013). Money and trust among strangers. *Proceedings of the National Academy of Sciences 110*(37), 14889–14893.

Charness, G. and C.-L. Yang (2014). Starting small toward voluntary formation of efficient large groups in public goods provision. *Journal of Economic Behavior & Organization 102*, 119–132.

Dal Bó, P. and G. R. Fréchette (2011, February). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review 101*(1), 411–29.

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

Dal Bó, P. and G. R. Fréchette (2019, November). Strategy choice in the infinitely repeated prisoner's dilemma. *American Economic Review 109*(11), 3929–52.

Duffy, J. and J. Ochs (2009). Cooperative behavior and the frequency of social interaction. *Games and Economic Behavior 66*(2), 785–812. Special Section In Honor of David Gale.

Ellison, G. (1994, 07). Cooperation in the Prisoner's Dilemma with Anonymous Random Matching. *The Review of Economic Studies 61*(3), 567–588.

Engl, F., A. Riedl, and R. Weber (2021, November). Spillover effects of institutions on cooperative behavior, preferences, and beliefs. *American Economic Journal: Microeconomics 13*(4), 261–99.

Fehr, E. and U. Fischbacher (2004). Social norms and human cooperation. *Trends in Cognitive Sciences 8*(4), 185–190.

Fehr, E. and I. Schurtenberger (2018). Normative foundations of human cooperation. *Nature human behaviour 2*(7), 458–468.

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

Friedman, J. W. (1971, 01). A Non-cooperative Equilibrium for Supergames12. *The Review of Economic Studies 38*(1), 1–12.

Fudenberg, D., D. G. Rand, and A. Dreber (2012, April). Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review 102*(2), 720–49.

Galbiati, R., E. Henry, and N. Jacquemet (2018). Dynamic effects of enforcement on cooperation. *Proceedings of the National Academy of Sciences 115*(49), 12425–12428.

Ghidoni, R., B. L. Cleave, and S. Suetens (2019). Perfect and imperfect strangers in social dilemmas. *European Economic Review 116*, 148–159.

Gneezy, U., A. Leibbrandt, and J. A. List (2016). Ode to the sea: Workplace organizations and norms of cooperation. *The Economic Journal 126*(595), 1856–1883.

Honhon, D. and K. Hyndman (2020). Flexibility and Reputation in Repeated Prisoners' Dilemma Games. *Management science* (September).

Huck, S., H.-T. Normann, and J. Oechssler (2004). Two are few and four are many: number effects in experimental oligopolies. *Journal of Economic Behavior Organization 53*(4), 435–446.

Isaac, R. M. and J. M. Walker (1988). Group size effects in public goods provision: The voluntary contributions mechanism. *The Quarterly Journal of Economics 103*(1), 179–199.

Kandori, M. (1992). Social Norms and Community En orcement. *59*(1), 63–80.

Molina, M., V. Nee, and H. Holm (2023). Cooperation with strangers: Spillover of community norms. *Organization Science 34*(6), 2315–2331.

Nosenzo, D., S. Quercia, and M. Sefton (2015, 3). Cooperation in small groups: the effect of group size. *Experimental Economics 18*, 4–14.

Ostrom, E. (2000). Collective action and the evolution of social norms. *The Journal of Economic Perspectives 14*(3), 137–158.

Peysakhovich, A. and D. G. Rand (2016). Habits of virtue: Creating norms of cooperation and defection in the laboratory. *Management Science 62*(3), 631–647.

Ranehill, E., F. Schneider, and R. A. Weber (2014, 3). Growing groups, cooperation, and the rate of entry. *SSRN Electronic Journal*.

Weber, R. A. (2006, March). Managing growth to achieve efficient coordination in large groups. *American Economic Review 96*(1), 114–126.

A Omitted regressions

A.1 The effect of "part" on cooperation, by treatment.

To test whether cooperation evolved significantly in the different parts of the experiment, we run logit regressions explaining cooperation by a categorical variable indicating the part of the experiment. We introduce random-effects at the individual level and demographic controls, and we use standard errors clustered at the highest group level to account for potential intra-group correlation: the group of eight in B8 and G, the group of four in B4, and the pair in B2. Table A.1 reports the marginal effects of eight regression models. Part 1 serves as the reference category; thus, each marginal effect should be interpreted as a difference relative to Part 1. Models (1) to (4) consider the individual choice to cooperate. Models (5) to (8) consider the occurrence of mutual cooperation. Each model restricts the estimation sample to a specific treatment, indicated in the line "Treatment" at the bottom of the table. At the bottom of Table A.1, we also report p-values from Wald tests comparing the marginal effects of Part 2 and Part 3.

	(1) Coop.=1	(2) Coop.=1	(3) Coop.=1	(4) Coop.=1	(5) Mut. Coop.=1	(6) Mut. Coop.=1	(7) Mut. Coop.=1	(8) Mut. Coop.=1
Part 1	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
Part 2	-0.0508**	-0.0880**	-0.204***	0.110***	-0.0382**	-0.0626*	-0.246***	0.173***
	(0.025)	(0.041)	(0.023)	(0.042)	(0.016)	(0.034)	(0.031)	(0.059)
Part 3	-0.0789***	-0.177***	-0.218***	0.141***	-0.0421***	-0.104***	-0.268***	0.220***
	(0.018)	(0.052)	(0.045)	(0.047)	(0.013)	(0.040)	(0.039)	(0.066)
N	13824	6776	15632	5169	6912	3388	7816	2519
Treatment	B8	B4	G	B2	B8	B4	G	B2
pvalue Part 2 v Part 3	.067	.029	.776	.333	.389	.147	.566	.333
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors clustered at the highest group level (the group of eight in B8 and G, the group of four in B4, and the pair in B2) are reported in parentheses. Marginal effect from logit models are reported. Models (1) to (4) include random effects at the individual level. Individual controls include age, gender, professional status, monthly expenses, difficulty understanding the instructions and past participation in behavioral experiments.

Table A.1: Cooperation across parts for the different treatments

B Strategy elicitation.

We estimate the distribution of individual strategies in the last part of the experiment using the Strategy Frequency Elicitation Method (see e.g. Dal Bó and Fréchette, 2011; Camera et al., 2012; Fudenberg et al., 2012; Bland, 2020). This method infers the most likely distribution of strategies from observed behavior, conditional on a predefined set of admissible strategies, by maximizing the likelihood of the data.

We estimate the prevalence of 4 strategies. Among these strategies, two are unconditional and correspond to clusters identified in the results section: Always cooperate and Always defect. The two remaining strategies are conditional: Grim-Trigger starts with cooperation and definitively switches to defection after encountering defection. Tit-for-tat starts with cooperation and then mimics the partner's previous play. These strategies are the most prevalent in past literature on stranger matching prisoner's dilemma (Camera et al., 2012).[9] We use the R package developed by Bland (2020) to recover the frequency of each strategy across treatments.

	Always Defect	Always Cooperate	Grim	TFT
B8	0.741***	0.039**	0.026	0.194***
	(0.041)	(0.017)	(0.021)	(0.038)
B4	0.632***	0.066**	0.058	0.244***
	(0.064)	(0.027)	(0.038)	(0.059)
G	0.534***	0.122***	0.169***	0.175***
	(0.045)	(0.027)	(0.041)	(0.035)
B2	0.168**	0.270**	0.260**	0.303
	(0.079)	(0.134)	(0.132)	(0.201)

Notes: ${}^{*}p < 0.10$, ${}^{**}p < 0.05$, ${}^{***}p < 0.01$. p-values from z-tests. Bootstrap standard errors in parentheses.

Table B.1: Distribution of strategies according to SFEM.

The results reported in Table B.1 are very consistent with what we found using cluster analysis. For instance, Always cooperate is estimated at 3.9% in Baseline 8, and 12.2% in Growth, and the difference is significant at the 1% level according to a z-test. Note, in addition, that Grim is estimated to be more common in Growth than in Baseline 8 or Baseline 4.

[9] Our main conclusions are unaffected by the inclusion of more or less strategies.

C Details on treatments and sessions.

Session num	Treatment	Nb Subjects	Nb Periods
1	Baseline 8	16	135
2	Growth	16	59
3	Growth	24	153
4	Baseline 8	24	101
5	Baseline 8	24	114
6	Growth	24	96
7	Growth	24	109
8	Baseline 8	24	89
9	Growth	24	143
10	Growth	24	65
11	Baseline 8	24	74
12	Growth	16	69
13	Baseline 8	16	81
14	Baseline 8	16	81
15	Baseline 2	22	79
16	Baseline 2	20	88
17	Baseline 4	20	102
18	Baseline 4	20	76
19	Baseline 4	16	51
20	Baseline 4	20	120
21	Baseline 2	14	125

Table C.1: Experimental Sessions and Treatment Details.

D Instructions (translated from French)

D.1 Instructions for treatment B8

Thank you for participating in this experiment on decision-making. You are not allowed to communicate with one another for the entire duration of the experiment. Your decisions will remain anonymous.

During the experiment, your earnings will be expressed in Experimental Monetary Units (EMU). At the end of the experiment, the EMUs you have earned will be converted into euros. The conversion rate is 100 EMU = €1.

The experiment is divided into 9 sequences. Each sequence consists of a certain number of rounds. In each round, you will take part in a game with a co-participant.

The instructions below describe the game, followed by details on each sequence. A diagram on the second sheet summarizes the structure of the experiment.

Your payment for the experiment will consist of the sum of the earnings accumulated across all sequences, plus a fixed payment of €7.50 for arriving on time to the experiment. Your payment will be transferred to you via Lydia, following the usual procedure of the Anthropo-lab. Please make sure to provide your phone number correctly when prompted.

The game

In the game, you will interact with a co-participant. The matching procedure that determines your co-participant in each round is explained in detail below.

In each round, you must choose one of two possible actions: A or B. The table below describes the decisions to be made during the game and the corresponding payoffs for you and your co-participant.

	Your co-participant chooses A	**Your co-participant chooses B**
You choose A	You: 20, Co-participant: 20	You: 0, Co-participant: 30
You choose B	You: 30, Co-participant: 0	You: 10, Co-participant: 10

The table reads as follows:

- If you choose A and your co-participant chooses A, you receive 20.

- If you choose A and your co-participant chooses B, you receive 0.
- If you choose B and your co-participant chooses A, you receive 30.
- If you choose B and your co-participant chooses B, you receive 10.

Note that your co-participant is in the same situation as you. This means that:

- If your co-participant chooses A and you choose A, (s)he receives 20.
- If your co-participant chooses A and you choose B, (s)he receives 0.
- If your co-participant chooses B and you choose A, (s)he receives 30.
- If your co-participant chooses B and you choose B, (s)he receives 10.

At the end of each round, you will be informed of your co-participant's choice as well as your payment for that round. A screenshot of the choice interface is provided on the second sheet.

Structure of the sequences

Before the start of the first sequence, you will be assigned to a group of 8 participants: yourself and 7 others. This group will remain fixed for the entire duration of the experiment, and you will only interact with members of your group throughout the experiment.

At the beginning of each round, you will be randomly paired with one of the 7 other members of your group, who will be your co-participant for that round. You will then play the game described above with this co-participant.

The number of rounds in each sequence is not known in advance. At the end of each round, it is randomly determined whether the sequence will end or continue for an additional round using the following procedure: The computer randomly draws a number between 1 and 100. If the number drawn is less than or equal to 90, the sequence continues for another round. Otherwise, the sequence ends. Thus, there is a 10% chance that the sequence will end after any given round.

If the current sequence is not the 9th and final sequence of the experiment, a new sequence will begin. If the current sequence is the 9th, the experiment ends.

Before the start of the first sequence, you will be asked to complete a short questionnaire to assess your understanding of the instructions. The experiment will begin afterward. After the 9th sequence, and before leaving the experiment, you will be asked to complete a brief demographic questionnaire.

D.2 Instructions for treatment B4

Thank you for participating in this experiment on decision-making. You are not allowed to communicate with one another for the entire duration of the experiment. Your decisions will remain anonymous.

During the experiment, your earnings will be expressed in Experimental Monetary Units (EMU). At the end of the experiment, the EMUs you have earned will be converted into euros. The conversion rate is 100 EMU = €1.

The experiment is divided into 9 sequences. Each sequence consists of a certain number of rounds. In each round, you will take part in a game with a co-participant.

The instructions below describe the game, followed by details on each sequence. A diagram on the second sheet summarizes the structure of the experiment.

Your payment for the experiment will consist of the sum of the earnings accumulated across all sequences, plus a fixed payment of €7.50 for arriving on time to the experiment. Your payment will be transferred to you via Lydia, following the usual procedure of the Anthropo-lab. Please make sure to provide your phone number correctly when prompted.

The game

In the game, you will interact with a co-participant. The matching procedure that determines your co-participant in each round is explained in detail below.

In each round, you must choose one of two possible actions: A or B. The table below describes the decisions to be made during the game and the corresponding payoffs for you and your co-participant.

	Your co-participant chooses A	**Your co-participant chooses B**
You choose A	You: 20, Co-participant: 20	You: 0, Co-participant: 30
You choose B	You: 30, Co-participant: 0	You: 10, Co-participant: 10

The table reads as follows:

- If you choose A and your co-participant chooses A, you receive 20.
- If you choose A and your co-participant chooses B, you receive 0.

- If you choose B and your co-participant chooses A, you receive 30.
- If you choose B and your co-participant chooses B, you receive 10.

Note that your co-participant is in the same situation as you. This means that:

- If your co-participant chooses A and you choose A, (s)he receives 20.
- If your co-participant chooses A and you choose B, (s)he receives 0.
- If your co-participant chooses B and you choose A, (s)he receives 30.
- If your co-participant chooses B and you choose B, (s)he receives 10.

At the end of each round, you will be informed of your co-participant's choice as well as your payment for that round. A screenshot of the choice interface is provided on the second sheet.

Structure of the sequences

Before the start of the first sequence, you will be assigned to a group of 4 participants: yourself and 3 others. This group will remain fixed for the entire duration of the experiment, and you will only interact with members of your group throughout the experiment.

At the beginning of each round, you will be randomly paired with one of the 3 other members of your group, who will be your co-participant for that round. You will then play the game described above with this co-participant.

The number of rounds in each sequence is not known in advance. At the end of each round, it is randomly determined whether the sequence will end or continue for an additional round using the following procedure: The computer randomly draws a number between 1 and 100. If the number drawn is less than or equal to 90, the sequence continues for another round. Otherwise, the sequence ends. Thus, there is a 10% chance that the sequence will end after any given round.

If the current sequence is not the 9th and final sequence of the experiment, a new sequence will begin. If the current sequence is the 9th, the experiment ends.

Before the start of the first sequence, you will be asked to complete a short questionnaire to assess your understanding of the instructions. The experiment will begin afterward. After the 9th sequence, and before leaving the experiment, you will be asked to complete a brief demographic questionnaire.

D.3 Instructions for treatment B2

Thank you for participating in this experiment on decision-making. You are not allowed to communicate with one another for the entire duration of the experiment. Your decisions will remain anonymous.

During the experiment, your earnings will be expressed in Experimental Monetary Units (EMU). At the end of the experiment, the EMUs you have earned will be converted into euros. The conversion rate is 100 EMU = €1.

The experiment is divided into 9 sequences. Each sequence consists of a certain number of rounds. In each round, you will take part in a game with a co-participant.

The instructions below describe the game, followed by details on each sequence. A diagram on the second sheet summarizes the structure of the experiment.

Your payment for the experiment will consist of the sum of the earnings accumulated across all sequences, plus a fixed payment of €7.50 for arriving on time to the experiment. Your payment will be transferred to you via Lydia, following the usual procedure of the Anthropo-lab. Please make sure to provide your phone number correctly when prompted.

The game

In the game, you will interact with a co-participant.

In each round, you must choose one of two possible actions: A or B. The table below describes the decisions to be made during the game and the corresponding payoffs for you and your co-participant.

	Your co-participant chooses A	**Your co-participant chooses B**
You choose A	You: 20, Co-participant: 20	You: 0, Co-participant: 30
You choose B	You: 30, Co-participant: 0	You: 10, Co-participant: 10

The table reads as follows:

- If you choose A and your co-participant chooses A, you receive 20.
- If you choose A and your co-participant chooses B, you receive 0.
- If you choose B and your co-participant chooses A, you receive 30.

- If you choose B and your co-participant chooses B, you receive 10.

Note that your co-participant is in the same situation as you. This means that:

- If your co-participant chooses A and you choose A, (s)he receives 20.
- If your co-participant chooses A and you choose B, (s)he receives 0.
- If your co-participant chooses B and you choose A, (s)he receives 30.
- If your co-participant chooses B and you choose B, (s)he receives 10.

At the end of each round, you will be informed of your co-participant's choice as well as your payment for that round. A screenshot of the choice interface is provided on the second sheet.

Structure of the sequences

Before the start of the first sequence, you will be assigned to a co-participant. This participant remains the same for duration of the experiment, and you will interact only with this co-participant for the length of the experiment. You will play the game described above with this co-participant.

The number of rounds in each sequence is not known in advance. At the end of each round, it is randomly determined whether the sequence will end or continue for an additional round using the following procedure: The computer randomly draws a number between 1 and 100. If the number drawn is less than or equal to 90, the sequence continues for another round. Otherwise, the sequence ends. Thus, there is a 10% chance that the sequence will end after any given round.

If the current sequence is not the 9th and final sequence of the experiment, a new sequence will begin. If the current sequence is the 9th, the experiment ends.

Before the start of the first sequence, you will be asked to complete a short questionnaire to assess your understanding of the instructions. The experiment will begin afterward. After the 9th sequence, and before leaving the experiment, you will be asked to complete a brief demographic questionnaire.

D.4 Instructions for treatment G

Thank you for participating in this experiment on decision-making. You are not allowed to communicate with one another for the entire duration of the experiment. Your decisions will remain anonymous.

During the experiment, your earnings will be expressed in Experimental Monetary Units (EMU). At the end of the experiment, the EMUs you have earned will be converted into euros. The conversion rate is 100 EMU = €1.

The experiment is divided into 9 sequences. Each sequence consists of a certain number of rounds. In each round, you will take part in a game with a co-participant.

The instructions below describe the game, followed by details on each sequence. A diagram on the second sheet summarizes the structure of the experiment.

Your payment for the experiment will consist of the sum of the earnings accumulated across all sequences, plus a fixed payment of €7.50 for arriving on time to the experiment. Your payment will be transferred to you via Lydia, following the usual procedure of the Anthropo-lab. Please make sure to provide your phone number correctly when prompted.

The game

In the game, you will interact with a co-participant. The matching procedure that determines your co-participant in each round is explained in detail below. The matching procedure differs across the various parts of the experiment.

In each round, you must choose one of two possible actions: A or B. The table below describes the decisions to be made during the game and the corresponding payoffs for you and your co-participant.

	Your co-participant chooses A	**Your co-participant chooses B**
You choose A	You: 20, Co-participant: 20	You: 0, Co-participant: 30
You choose B	You: 30, Co-participant: 0	You: 10, Co-participant: 10

The table reads as follows:

- If you choose A and your co-participant chooses A, you receive 20.
- If you choose A and your co-participant chooses B, you receive 0.

- If you choose B and your co-participant chooses A, you receive 30.
- If you choose B and your co-participant chooses B, you receive 10.

Note that your co-participant is in the same situation as you. This means that:

- If your co-participant chooses A and you choose A, (s)he receives 20.
- If your co-participant chooses A and you choose B, (s)he receives 0.
- If your co-participant chooses B and you choose A, (s)he receives 30.
- If your co-participant chooses B and you choose B, (s)he receives 10.

At the end of each round, you will be informed of your co-participant's choice as well as your payment for that round. A screenshot of the choice interface is provided on the second sheet.

Part 1

At the beginning of Part 1, each participant is randomly assigned to a specific co-participant. You will interact with the same co-participant throughout the entire Part 1. However, note that all interactions are anonymous. Part 1 consists of 3 sequences. Each sequence comprises a number of rounds, which is determined according to the procedure described in the following paragraph. In each round, you will play the game described above with your assigned co-participant. The number of rounds in each sequence is not known in advance. At the end of each round, it is randomly determined whether the sequence continues for another round or ends, according to the following procedure: The computer draws a random number between 1 and 100. If the number is less than or equal to 90, the sequence continues for another round. Otherwise, the sequence ends. Thus, there is a 10% chance that the sequence will end after any given round. If the current sequence is not the third and final sequence of Part 1, a new sequence begins with the same co-participant. If the current sequence is the third sequence of Part 1, then Part 1 ends and Part 2 begins.

Part 2

Like Part 1, Part 2 also consists of 3 sequences during which you play the same game over a number of rounds. As in Part 1, the end of each sequence is determined randomly, with a 10%

chance of stopping after each round. The only difference from Part 1 is that you no longer play with the same co-participant in every round. In Part 2, you interact within a group of 4 participants. This group consists of the two members of your Part 1 group (you and your Part 1 co-participant) and two members from another randomly selected Part 1 group. At the beginning of each round, your co-participant is randomly selected from the 3 other members of your Part 2 group. Each of these 3 members has an equal probability of being selected as your co-participant.

Part 3

Like Parts 1 and 2, Part 3 also consists of 3 sequences in which you once again play the game described above over a number of rounds. As in the previous parts, the end of each sequence is randomly determined, with a 10% chance of stopping after each round. The only difference from the previous parts is that in Part 3, you interact within a group of 8 participants. This group includes the 4 members of your Part 2 group (you and the other 3 members), and 4 members from another Part 2 group that is randomly selected. At the beginning of each round, your co-participant is randomly selected from the 7 other members of your Part 3 group. Each of the 7 members has an equal probability of being selected as your co-participant.
Before the start of the first sequence, you will be asked to complete a short questionnaire to assess your understanding of the instructions. The experiment will begin afterward. After the 9th sequence, and before leaving the experiment, you will be asked to complete a brief demographic questionnaire.

D.5 Diagram summarizing the structure of the experiment

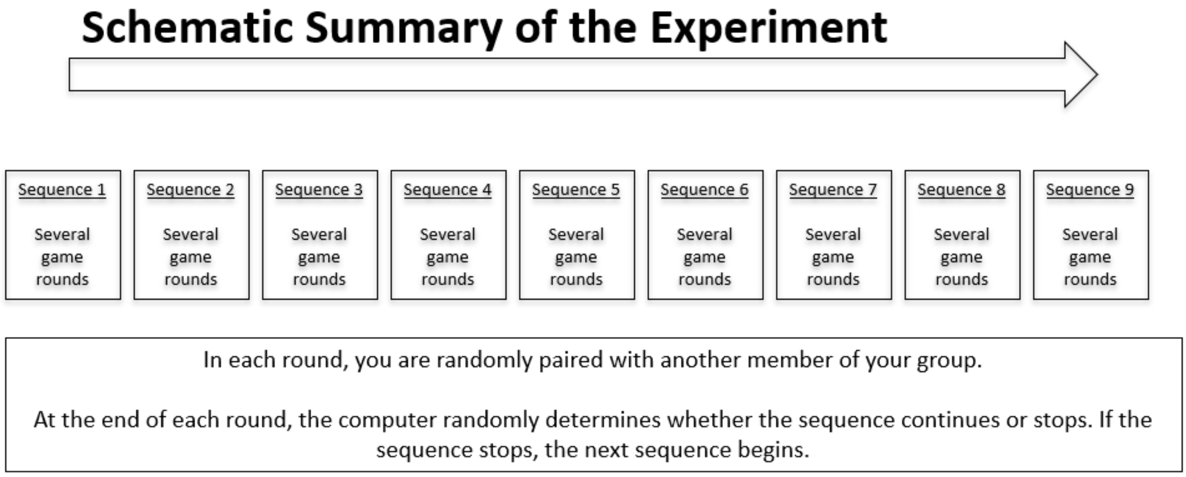

Figure D.1: Diagram summarizing the structure of the experiment, appended to the instructions.

D.6 Screenshot

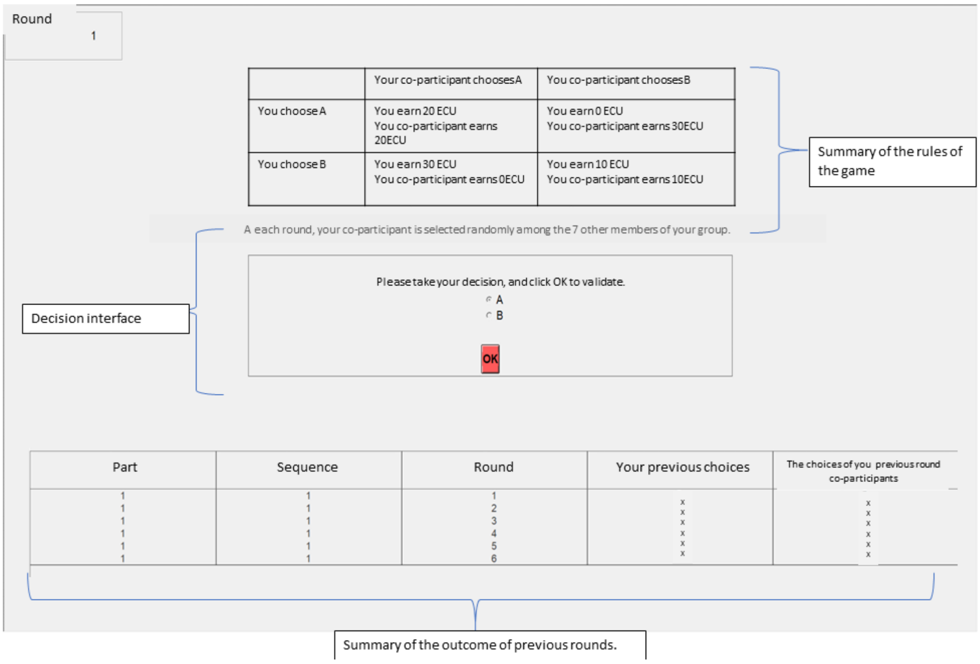

Figure D.2: Screenshot of the decision interface with annotations, appended to the instructions.