Voluntary participation and optimism: How and when outside individual options can facilitate group collaboration

Ryutaro Mori  
The University of Tokyo; Japan Society for the Promotion of Science  https://orcid.org/0000-0001-9131-4184

Nobuyuki Hanaki  
Osaka University  https://orcid.org/0000-0003-3127-690X

Tatsuya Kameda  
Tatsuya Kameda (tkameda@l.u-tokyo.ac.jp)  The University of Tokyo; Hokkaido University; Tamagawa University  https://orcid.org/0000-0002-6666-493X

Research Article

Keywords: collaboration, voluntary participation, experimental game

Posted Date: October 5th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3300738/v3

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Voluntary participation and optimism: How and when outside individual options can facilitate group collaboration

Ryutaro Mori
Department of Social Psychology, The University of Tokyo, Japan
Japan Society for the Promotion of Science, Japan

Nobuyuki Hanaki
Institute of Social and Economic Research, Osaka University, Japan

Tatsuya Kameda
Department of Social Psychology, The University of Tokyo, Japan
Center for Experimental Research in Social Sciences, Hokkaido University, Japan
Brain Science Institute, Tamagawa University, Japan

Whether in business ventures or academic research, voluntary participation is at the heart of collaboration. Yet, previous studies concerning collaboration have primarily focused on the free-riding problem, largely overlooking this mundane but critical aspect. We address this gap by modeling collaborative situations as people’s voluntary choices between joining an uncertain public goods provisioning in groups and pursuing a certain but less profitable individual option. Theoretically and experimentally, we show that this voluntary nature in collaboration increases the likelihood of group success by filtering out otherwise pessimistic defectors or encouraging them toward cooperation. Further, we reconcile our findings with the existing literature highlighting detrimental effects of individual alternatives to collaborative solutions. We argue that whether individual options aid or hinder collaboration hinges on the degree of the “externality” of loners (who pursue the individual option) to the group. As long as the collaboration pertains to local public goods, the externality remains limited, ensuring that the existence of an individual option robustly aids collaborative success by fostering cooperation through improved optimism within groups.

INTRODUCTION

Ranging from business ventures to academic research, collaboration is a robust strategy in human societies to achieve objectives that can never be accomplished by any individual working alone\(^1\)\(^-\)\(^3\). However, initiating successful collaborations is not at all trivial, as it requires costly and coordinated efforts from multiple individuals\(^3\)\(^-\)\(^6\). Specifically, each individual has an incentive to exploit the collective output without incurring personal effort, discouraging many from committing to collaboration. Presumably because the major difficulty stems from this free rider problem, collaboration has been mainly modeled as public goods provisioning. There, individuals are bound within a group with fixed memberships and must decide whether to cooperate with the group. Several mechanisms that mitigate the free rider problem have been identified, including other-regarding preferences\(^7\)\(^-\)\(^9\), internalized norms\(^10\)\(^,11\), peer punishment\(^12\)\(^-\)\(^15\), and reciprocity in repeated interactions\(^16\)\(^-\)\(^20\).

Yet, such models have largely sidelined the fact that many collaborations in the real world do not involve the entire public or take place within predetermined group boundaries: People often have individual alternatives outside groups and are thus free to opt in or out of group endeavors. In the case of a startup company, for instance, only those who voluntarily choose to join will work together. Consequently, even though the company still faces the issue of free riding, those who opt not to participate in the collaborative venture remain uninvolved.

To understand group collaboration in such voluntary situations, we extend previous models of public goods provision and coordination problems using threshold public goods game\(^21\)\(^-\)\(^30\) (Fig. 1A lower left). In a group, each member chooses either to contribute their endowment to the group or to keep it for themselves. Only if a sufficient number (“threshold”) of individuals have cooperated, the group succeeds to produce the public goods which are to be shared equally among its members. Crucially, we contrast this uncertain but potentially rewarding group collaboration with a certain but less profitable individual option outside the group: Players are free to pursue an individual option outside of groups or to participate in group collaboration with others who have also voluntarily opted in (Fig. 1A).

Recently, an influential line of research\(^26\)\(^-\)\(^28,31\) has employed a similar (but distinct, as will be discussed later) setting, and has experimentally demonstrated that the
presence of individual options can have detrimental effects on group collaboration. They argued that just as billionaires may not be highly motivated to contribute their share to the social insurance system, the ability of self-reliance crowds out collaborative motivations, and groups may have to struggle even more to get enough contribution.

Here we challenge this view, arguing that the availability of an individual option renders collaboration voluntary and thereby facilitates, rather than hinders, successful group collaboration. We first develop theoretical predictions about the difference in collaborative success under the presence or absence of the individual option (i.e., voluntary vs. mandatory participation in the group). If the number of cooperators is equal to or more than the predetermined threshold value, \( q (q = 2, 4, \text{or } 5) \), then all the group members earn 30 extra points. When choosing the individual option, players are guaranteed a smaller additional payoff of 10 points irrespective of other players’ choices.

Further, two distinct behavioral mechanisms—the self-selection of optimists into groups \( ^{12} \) and some pessimists switching from defection to cooperation by becoming optimistic—were responsible for the increase in cooperation rates.

Finally, we argue that an apparent contradiction between the positive (current study) and the negative \( ^{26–28} \) effects of individual alternatives to collaboration can be synthesized via the concept of loners’ externality to the collective outcome: the varying degrees of impact that players who leave the group (or choose an individual option) still exert on group endeavors. We show that when the collaboration concerns public goods for the entire population or a fixed group, the large externality of loners can deteriorate collaboration, whereas when the collaboration concerns local public goods with flexible members under voluntary participation, an individual option facilitates collaboration with improved optimism.

**RESULTS**

Modeling collaboration using a voluntary threshold public goods game

---

Figure 1. Modeling collaboration under voluntary participation. (A) Individuals can either opt in to group collaboration or choose an individual option outside of the group. When opting to collaborate, players are randomly assigned to groups of five members (if the total number of players opting in is not a multiple of five, any remaining players are assigned to the individual option). Group members can choose whether to cooperate (incurring the private cost of 10 points) on the group project. If the number of cooperators is equal to or more than the predetermined threshold value, \( q (q = 2, 4, \text{or } 5) \), then all the group members earn 30 extra points. When choosing the individual option, players are guaranteed a smaller additional payoff of 10 points irrespective of other players’ choices. (B, C) Expected payoff (B) and best response (C) as a function of subjective belief about how likely others are to cooperate within groups (i.e., \( \gamma \)). Each color represents a possible action under mandatory or voluntary participation: orange = cooperate (C), blue = defect (D), and green = leave (L). Recall that because leave is not available under mandatory participation, expected payoffs in mandatory conditions can be compared only between cooperate and defect.
We used the following game for both theoretical development and an experiment. The game proceeds as summarized in Fig. 1A. All players start with the endowment of 10 points. They first decide whether to participate in a group-based collaboration or to employ an individual option outside of the group (“leave”). Choosing the individual option secures 10 extra points irrespective of other players’ decisions, plus the initial endowment (10 points). When choosing to participate in collaboration, players are randomly assigned to groups of five members and play the threshold public goods game (in the experiment, if the total number of participants opting for groups is not a multiple of five, the remaining participants are assigned to the individual option). In groups, each member chooses whether to invest their initial endowment of 10 points in the group (“cooperate”) or to keep it for themselves (“defect”). If the number of cooperators fails to gather enough cooperators, the collaboration does not create any gains, and the investments are not returned. Players are commonly informed of these rules at the outset, including the threshold value for group collaboration.

Critically, the presence/absence of the individual option corresponds to voluntary/mandatory participation in collaboration, respectively. Additionally, the threshold value determines the minimum number of cooperators for successful collaboration and thereby alters the degree of free-riding incentive within groups. With a higher threshold, collaboration requires more cooperators and creates less temptation to free ride, as we see below. Here, we considered three threshold values: \( q = 2, 4, \) or 5 (out of 5), representing a strong, weak, and null temptation to free ride, respectively. By manipulating these two factors orthogonally, we study how voluntary participation affects collaborative efforts under varying degrees of free-riding incentives within the collaboration.

**Theoretical analysis: whether/how voluntary participation can aid collaboration**

In the threshold public goods game, the action that maximizes the player’s own payoff depends on other players’ actions. Thus, we start by assuming that a player forms a subjective belief about others’ average cooperativeness (i.e., how likely it is that other group members will cooperate in a given situation) and selects their own action based on the belief. We denote the player’s belief about others’ cooperativeness as \( \gamma \in [0, 1] \), their action as \( x \in \chi \), and the resultant payoff as \( \pi \). When collaboration is mandatory, the set of possible actions for the player, \( \chi \), is \{C (cooperate), D (defect)\}, and when collaboration is voluntary, it is \{C, D, L (leave)\}.

Given their belief about others, the expected payoff for each action \( E[\pi|x = i] \) becomes:

\[
E[\pi|x = C] = 30 \times \Gamma_q - \gamma,
\]
\[
E[\pi|x = D] = 10 + 30 \times \Gamma_q,
\]
\[
E[\pi|x = L] = 10 + 10
\]

where \( \Gamma_q \) denotes the probability that at least \( q \) out of four other members will cooperate, according to the binomial expansion with their belief \( \gamma : \Gamma_q = \sum_{i=0}^{q} \binom{4}{i} \gamma^i (1 - \gamma)^{4-i} \). Figure 1B displays expected payoffs as a function of \( \gamma \) under three threshold values (\( q = 2, 4, \) or 5 from the left). Figure 1C further depicts the best responses (i.e., actions maximizing expected payoff given belief) under mandatory or voluntary participation.

Two observations are noteworthy. First, when participation is mandatory, defection arises as the best response in two separate regions (see blue intervals in Fig. 1C: Mandatory, left and middle). These regions may correspond to the operation of two different psychological motives. One is “fear” that cooperation may fail short even if the player themselves cooperates, and it operates when the player expects little cooperation from others (i.e., large \( \gamma \)). The other can be termed “greed”: When a player expects a great deal of cooperation from others (i.e., large \( \gamma \)), they may greedily defect to free ride on successful collaboration. Note that with the threshold of 5, greedy defection cannot exist, as free riding is no longer possible (i.e., a player’s own defection guarantees failed group endeavor).

Second, when participation is voluntary, the configuration of the best response changes drastically. As seen in the green intervals in Fig. 1C: Voluntary, leaving is the best possible action across the thresholds unless the player’s \( \gamma \) is sufficiently high. In other words, those who hold low expectations about their peers’ cooperation, including those who would defect out of fear under mandatory participation, will naturally choose the individual option outside the group. Hence, voluntary participation can work as a self-selection mechanism that filters pessimistic defectors out of collaborative efforts.

To reasonably predict the population-level cooperation rate resulting from the individual best-response function (Fig. 1C), we further need to consider the distribution of subjective beliefs in the population, \( \phi(\gamma) \). The proportion of the population that chooses action \( x, r_x \), can then be obtained by integrating the best response with respect to \( \gamma \):

\[
r_x = \int_0^1 I\{\text{Best response is } x, \text{given } \gamma\} \phi(\gamma) \, d\gamma, \tag{2}
\]

where \( I\{A\} \) is the indicator function (returns 1 if \( A \) is true and 0 otherwise). Accordingly, the cooperation rate in the population, \( p_{\text{coop}} \), is obtained as follows:

\[
p_{\text{coop}} = \frac{r_C}{r_C + r_D}. \tag{3}
\]

Let us now compare the population-level cooperation rate between voluntary and mandatory participation. Consider first when the threshold value is 5. Here, if participation is voluntary, there is no interval of \( \gamma \) where the best response is defection (see Fig. 1C right), which leads to \( r_D = 0 \) irrespective of the distribution of beliefs.
Therefore, the resultant cooperation rate, \( p_{\text{coop}} \), equals 1 under voluntary participation (as long as \( r_c > 0 \)). In contrast, with smaller threshold values (see Fig. 1C left and middle), where there remain possibilities for free riding even under voluntary participation, the results should depend on \( \phi(\gamma) \). We investigated how \( p_{\text{coop}} \) differs between mandatory and voluntary participation under various specifications of \( \phi(\gamma) \). Without losing much generality, we can model \( \phi(\gamma) \) with a Beta distribution and calculate \( p_{\text{coop}} \) while changing its parameters systematically. We found that the resultant \( p_{\text{coop}} \) is higher under voluntary participation than under mandatory participation across a wide range of parameters (see Fig. S1 in the Supplementary Materials). Overall, we can predict that voluntary participation will induce the self-selection of optimists into groups and of pessimists into individual options, which consequentially increases the population-level cooperation rate for group endeavors.

We have considered how voluntary participation may alter the member composition of group endeavors via self-selection dynamics. Notice that, thus far, we have assumed that the distribution of players’ beliefs is fixed, whether participation is mandatory or voluntary. In other words, we have argued that players may choose different actions in these two situations (e.g., defecting vs. leaving according to the best-response function in Fig. 1C) but keep their subjective beliefs \( \gamma \) intact. However, some players may engage in reasonings similar to ours and further update their beliefs about others’ cooperativeness, which potentially increases (or decreases) cooperation under voluntary participation. Indeed, when the threshold is 5, our analysis above suggests that players can reasonably predict that no one will defect under voluntary participation, because players are always better off leaving than defecting. Thus, one can argue that in the voluntary condition, players come to expect 100% cooperation from other members who participate in groups and then decide themselves to behave cooperatively.

With smaller threshold values (Fig. 1C left and middle), the predictions, again, become less clear because of the possibility of greedy defection. At first, players may take the self-selection dynamics into account, forming more optimistic beliefs. However, they may further update their subjective beliefs recursively, fluctuating between optimism and pessimism. For instance, a player who initially anticipates increased cooperation via the self-selection may end up with pessimism by reasoning that other players will have the same anticipation and begin free riding with the updated beliefs.

If we consider only the equilibrium case where players’ beliefs and actual plays converge under best responses, voluntary participation is predicted to induce greater cooperation than mandatory participation (in a nutshell, because a noncooperative equilibrium of groups where no one cooperates yields lower return than the outside individual option, it is unsustainable and eliminated under voluntary participation; see the Supplementary Materials for details). Nevertheless, it is unclear a priori whether, or to what extent, people change their subjective expectations between mandatory and voluntary participation, especially under the smaller threshold values (\( q = 2, 4 \)). We thus leave this as an empirical question to be addressed with our experimental data.

To sum up, our theoretical analysis suggests two related but distinct mechanisms by which voluntary participation may facilitate collaboration. First, as a baseline, it will make optimists self-select into groups and pessimists into the individual option, increasing the cooperation rate within groups. The second possible mechanism is the updating of beliefs: Anticipating self-selection, individuals may further update their subjective beliefs and possibly become more optimistic about successful collaboration. Together, we conjecture that voluntary participation works as a natural device to filter pessimistic individuals (who would otherwise defect from fear) out of groups and/or encourage them to choose cooperation.

**Experimental tests of the predictions**

To test the predictions, we ran a preregistered behavioral experiment (\( N = 191 \)). We had around 30 (30 to 35) participants in each experimental session. In the session, we manipulated (1) whether players had the individual option (i.e., voluntary or mandatory participation) and (2) the threshold value (2, 4, or 5) of the public goods game, resulting in a 2 \( \times \) 3 factorial within-subject design. Participants played each of the six different conditions once in a randomized order with no feedback until the end of the whole experiment (see Methods for details). Within each condition, participants first estimated how likely others would be to choose each action and then selected their own actions in the game. Participants’ own action selections and estimations about others’ actions were both incentivized (see Methods for details). Additionally, we elicited participants’ economic and psychological characteristics that may partially account for heterogeneities in their play, including risk preference\(^{34}\) and inequity aversion\(^{35}\) (see Methods for details).

**Voluntary participation improved success rates and efficiencies of group collaboration**

Responding to the difference in the threshold values and whether participation was mandatory or voluntary, participants indeed changed their action selections (see Fig. S2 for the exact number of participants choosing each action in each condition). Figure 2A displays rates of participants’ cooperation for the public good within groups in each condition. Participants cooperated significantly more under the voluntary conditions compared to the mandatory conditions across the three thresholds (the difference in cooperation rate; \( \Delta p_{\text{coop}} \): \( q = 2 \): \( \Delta p_{\text{coop}} = 0.16 \), bootstrap 95% confidence interval [CI]; hereafter, all CIs refer to bootstrapped 95% CI unless...
otherwise noted; see Methods for the details of the bootstrapping procedure] [0.06, 0.25]; \( q = 4 \): \( \Delta p_{\text{coop}} = 0.22 \), CI [0.14, 0.31]; \( q = 5 \): \( \Delta p_{\text{coop}} = 0.21 \), CI [0.16, 0.27]). Voluntary participation elevated the cooperation rates by around 20 points for each threshold.

To evaluate how the increase in individual cooperation rate affected the outcomes of groups, we next looked at population-level statistics. One important measure is how likely groups are to succeed in generating the collective benefits (“group success rate”; probability of a group gathering enough cooperators). As seen in Fig. 2B, under the mandatory conditions, the group success rate decreased as the threshold got higher; only 29.9% of groups would succeed with the highest threshold of 5. In contrast, in the voluntary conditions, the group success rate reached nearly 1 across the thresholds. For each threshold, the group success rates were significantly higher in the voluntary conditions compared to the mandatory conditions, with the gaps being particularly prominent under higher threshold values (the difference in group success rate (\( \Delta p_{\text{success}} \)):

- \( q = 2 \): \( \Delta p_{\text{success}} = 0.12 \), CI [0.05, 0.20];
- \( q = 4 \): \( \Delta p_{\text{success}} = 0.41 \), CI [0.27, 0.56];
- \( q = 5 \): \( \Delta p_{\text{success}} = 0.70 \), CI [0.57, 0.80]).

Another key population-level metric is the average payoff, or efficiency. A rise in cooperation rates within the voluntary groups does not necessarily result in greater efficiency for two distinct reasons: overcooperation within groups and overpresence of loners outside of groups. Overcooperation could reduce the average payoff because contributions above the threshold do not generate any further gain for the group. Also, if there are too many loners whose payoffs (20 points) are lower than those of successful group members (30 points for cooperators and 40 points for defectors), that could also cause the average payoff to drop at the population level.

Figure 2C displays the average payoff normalized as a ratio against the most efficient net payoff possible in a group (i.e., with the exact threshold number of cooperators). Here, we distinguished between two efficiencies: the average payoff within groups while excluding loners and the average payoff for the entire population with loners included. The two differ under voluntary participation (plain or hatched cyan bars), whereas coincide (red bars) under mandatory participation by definition (as there are no loners). Notice that the potential issue of overcooperation can be assessed by comparing voluntary and mandatory participation in terms of the average payoff within groups, and the issue of too many loners by the average payoff for the entire population. Results indicate that in the former comparison, the average payoff was higher in the voluntary conditions than in the mandatory conditions across the thresholds (the difference in efficiency within groups: \( q = 2 \): 0.06, CI [0.02, 0.10]; \( q = 4 \): 0.32, 95% CI [0.20, 0.43]; \( q = 5 \): 0.63, CI [0.52, 0.71]). In the second comparison, the average payoff for the entire population was also higher in voluntary conditions except for \( q = 2 \) (the difference in efficiency including loners: \( q = 2 \): −0.07, CI [−0.11, −0.01]; \( q = 4 \): 0.18, CI [0.07, 0.28]; \( q = 5 \): 0.53, CI [0.43, 0.60]). For \( q = 2 \), more than 80% of the mandatory groups were already successful (see Fig. 2B left), yielding a much higher average payoff to their members (30 or 40 points) than that of loners (20 points). Taken together, the results indicate that voluntary participation did improve the efficiency among group members while avoiding overcooperation issues and could also improve the efficiency of the entire population against the issue of too many loners (when \( q = 4 \) or 5).
Two mechanisms underlying the collaborative success under voluntary participation

Having established that voluntary participation leads to higher cooperation rates and better collective outcomes than mandatory participation, we now turn to its mechanisms. Recall that our theoretical analysis suggested two possible pathways: Pessimistic players who would defect out of fear under mandatory participation will be filtered out from groups (i.e., will opt for the individual option) and/or be encouraged to form optimistic beliefs and turn to cooperation within groups.

To address the first point, we examined how participants' defection rates in the mandatory condition and leaving rates in the voluntary condition both correlate with their "original" beliefs about others’ cooperativeness ($\gamma$ in the mandatory condition). Figure 3A shows that under each threshold, more pessimistic participants (i.e., holding a smaller $\gamma$) were more likely to defect (blue) when placed in the mandatory condition (mixed-effect logistic regression: $\beta = -7.74$, CI $[-25.04, -8.71]$; see Methods for details) and to leave (green) groups for the individual option in the voluntary condition (mixed-effect logistic regression: $\beta = -1.73$, CI $[-2.81, -0.97]$). As seen in Fig. 3B, this means that defectors (darker green) in the mandatory condition were more likely to leave groups compared to cooperators (lighter green), under voluntary participation. (C) Nonlone’s cooperation rates in the two conditions (red: mandatory, cyan: voluntary). For the calculation of cooperation rates in the mandatory condition, we included only those participants who stayed in groups in the voluntary condition of the same threshold (i.e., nonlones), thereby eliminating the effects of self-selection. (D) Scatter plots of participants’ subjective beliefs about others’ cooperativeness ($\gamma$) in the mandatory (x axis) and voluntary (y axis) conditions for each threshold. Dot colors correspond to the changes in action (orange: positive change where a player changed action from defecting in the mandatory condition to cooperating in the voluntary condition, gray: no change, blue: negative change to defection). As in (C), we used solely the data from nonlones.

Figure 3. Voluntary participation filters out or encourages pessimistic defectors. (A) Rates of participants who defected in the mandatory condition (blue) and those who chose to leave groups in the voluntary condition (green), as a function of their original beliefs about others’ cooperativeness ($\gamma$ in the mandatory conditions; grouped in 0.1 increments). (B) Proportions of participants who left groups in the voluntary condition as a function of their action in the mandatory condition with the same threshold values. Across the thresholds, defectors in the mandatory condition (darker green) were more likely to leave groups compared to cooperators (lighter green), under voluntary participation. (C) Nonlone’s cooperation rates in the two conditions (red: mandatory, cyan: voluntary). For the calculation of cooperation rates in the mandatory condition, we included only those participants who stayed in groups in the voluntary condition of the same threshold (i.e., nonlones), thereby eliminating the effects of self-selection. (D) Scatter plots of participants’ subjective beliefs about others’ cooperativeness ($\gamma$) in the mandatory (x axis) and voluntary (y axis) conditions for each threshold. Dot colors correspond to the changes in action (orange: positive change where a player changed action from defecting in the mandatory condition to cooperating in the voluntary condition, gray: no change, blue: negative change to defection). As in (C), we used solely the data from nonlones.
(lighter green) in the mandatory condition (mixed-effect logistic regression: $\beta = 1.11$, CI [0.56, 1.69]). Hence, voluntary participation indeed worked as a self-selection mechanism via subjective beliefs, filtering pessimistic defectors out of collaborative groups. (Supplementary texts and Table S1 in the Supplementary Materials provide additional exploratory analyses concerning the relationships between participants’ economic and psychological characteristics and their individualistic choices in the threshold public goods game.)

Next, we consider the second pathway. Can voluntary participation encourage pessimists to become optimistic cooperators rather than merely prompting them to opt out from groups? To examine this issue, we focused on participants who stayed in groups under voluntary participation (i.e., excluding the loners) and checked how their cooperation rates differed between the mandatory and voluntary conditions. Even in this comparison, with threshold values of 4 and 5, the cooperation rates were significantly higher in the voluntary condition than in the mandatory condition (Fig. 3C; the difference in cooperation rates among nonloners: $q = 2 : 0.08$, CI $[-0.02, 0.19]$; $q = 4 : 0.09$, CI $[0.01, 0.18]$; $q = 5 : 0.14$, CI $[0.09, 0.20]$). Then, did this shift in cooperative action parallel the optimistic updating of beliefs as we conjectured? Indeed, the overall distributions of subjective beliefs indicate that participants became more optimistic in the voluntary condition than in the mandatory condition when the threshold value was 4 or 5 (paired-samples $t$ test: $q = 2 : t(187) = 1.23$, $p = .22$; $q = 4 : t(181) = 2.63$, $p = .01$; $q = 5 : t(180) = 5.97$, $p = 1.2e^{-8}$). Moreover, these optimistic belief changes coincided with cooperative behavioral changes, at the individual level. Figure 3D depicts participants’ subjective beliefs about others’ cooperativeness in the mandatory (x axis) and voluntary (y axis) conditions, with the color of dots indicating the direction of their action changes (see the caption of Fig. 3D for details). For each threshold value, players who switched from defection in the mandatory condition to cooperation in the voluntary condition (orange) were primarily distributed above the diagonal (i.e., had more optimistic beliefs in the voluntary condition). We indeed found a positive relationship between the change in action and the change in belief (a mixed-effect regression: $\beta = 0.82$, CI $[0.67, 1.04]$).

Together, these results suggest that when incentives for free riding are relatively small ($q = 4, 5$; Fig. 1B), voluntary participation encourages individuals to develop optimism regarding others’ cooperativeness and thereby promotes their own cooperative behavior. It is also important to notice that across the six conditions, participants’ defection rates decreased almost monotonically as their expectations about others’ cooperativeness increased (blue lines in Fig. 3A for the mandatory conditions; see Fig. S3 for the voluntary conditions). Whereas standard theories of expected utility maximization dictate that players should respond to the probability of their own decisions being pivotal (i.e., both necessary and sufficient for the provision of collective benefits$^{25}$), in our experimental data, raw expectations (i.e., “how likely are others to cooperate?”) were more predictive of their actions (whether to cooperate or defect) than the pivotal probabilities calculated from them across conditions (see Methods and the Supplementary Materials for the analysis of receiver operating characteristics). These results underscore the importance of fear, not greed, as the primary driver of defection within groups attempting to initiate collaboration under social uncertainty$^{24,33,36}$. Consequently, for successful collaboration, it is more critical to mitigate pessimistic defection than to prevent greedy defection. Voluntary participation just does that.

**Theoretical mapping: Reconciling seemingly contradictory effects of individual options**

With the theoretical and experimental analyses, we have demonstrated the positive impacts of voluntary participation (i.e., the presence of an outside individual option) on collaborative success. However, as we mentioned earlier, an influential line of research has recently highlighted the negative effects of individual options on coordination within groups$^{26,28,31}$. How can we reconcile the two seemingly conflicting results? Here, we propose an integrative view highlighting the critical difference between “global” and “local” public goods. Gross and colleagues$^{26}$ confronted participants with a variant of the threshold public goods game called the collective risk social dilemma$^{37}$. In their game, participants are first embedded to groups of four or five members (“village”) and face a shared risk of flooding to their “village” that causes the loss of all properties. Participants can choose either (a) investing their endowment to build a public dam that surrounds the entire village, (b) investing to build a personal dam only around their house, or (c) keeping their endowment to themselves without investing in any dam. By manipulating the cost–benefit ratio associated with constructing the dam collectively (Option a) versus individually (Option b), they examined how participants shift from collective to individual options. Crucially, at the intermediate levels of the cost–benefit ratio, where the uncertain collective solution can add only relatively little efficiency over the certain individual option, participants struggled to balance between self-reliance and interdependence: Some started to employ the individual solution while others continued to attempt to solve the problem collectively. Notice that, in their set-up, with more loners (i.e., players who choose the individual solution), a smaller number of villagers must share the cost to build the public dam that surrounds the entire village including the loners’ houses; consequently, in terms of provision of the public dam, loners function essentially the same as defectors (i.e., players who retain their endowment and do not fund any dam) in the village. In other words, since group boundaries are fixed prior to the individuals’ decision to opt for a collective or individual solution, the cooperation...
rate determining the collective outcome is the proportion of cooperators among the entire population in the village (fixed group). We call this situation “global public goods.”

In contrast, in our scenario focusing on collaboration\textsuperscript{1,2}, the cooperation rate determining the collective outcome for groups is the proportion of cooperators among players who opt in to groups. This is the natural consequence of our assumption that participation in collaboration is voluntary rather than mandatory. Groups are not fixed from the outset, and their formation comes after the individuals’ decision to opt in or not; groups consist only of those who join voluntarily. Note that this does not imply that people no longer suffer from the free-riding problem: Groups still must create shared benefits from the costly efforts of some members. We call this situation “local public goods”\textsuperscript{38–41}.

To facilitate a finer comparison, we introduce a new parameter $\rho \in [0, 1]$ reflecting the degree of loners’ impact (externality) on the collective outcome, as illustrated in Fig. 4A. The cooperation rate that ultimately determines whether the group (e.g., the entire population as in (26)). Right: Creating local public goods. Groups consist only of individuals who voluntarily opt in, and thus loners have no externality to the collective outcome. (B) Group success rate (computed from participants’ choice data, $r_C$, $r_D$, and $r_L$, in our experiment) as a function of the degree of loners’ externality, $\rho$. Blue lines correspond to the situation with the individual option (voluntary participation) and red to the situation without the individual option (mandatory participation).

Figure 4. Synthesizing positive and negative effects of the individual option on collaborative success via the degree of loners’ externality (impact). (A) Left: Creating global public goods. Groups are fixed in advance (or the entire population is one group); thus, loners have full externality on the collective outcome and effectively function the same as defectors to the group (e.g., the “village” example in (26)). Right: Creating local public goods. Groups consist only of individuals who voluntarily opt in, and thus loners have no externality to the collective outcome.

\[ p_{\text{coop}}(\rho) = \frac{r_C}{r_C + r_D + \rho r_L}. \] (4)

Assuming the group size is 5, the group success rate is computed as
where \( q \) is the threshold value.

The two scenarios presented above can be seen as the opposite extremes of the continuum. When \( \rho = 1 \), loners have full externality to the group outcome and function identically to defectors (Fig. 4A left). The cooperation rate determining the collective outcome is \( r_c = r^{\rho}_c + r^{p}_d + r^{l}_l \). This should be the case when collaborations are intended to produce global public goods, as in the village example. If collaboration necessarily involves all individuals, groups are fixed at the outset, and what matters to their outcome is the proportion of cooperators among the entire population, including loners.

In contrast, when \( \rho = 0 \), loners are separated from collaboration and thus exert no externality (Fig. 4A right). Consequently, the collective outcome will be determined by the ratio of cooperators to individuals who opt in to collaboration: \( \frac{r^{p}_c}{r^{\rho}_c + r^{p}_d + r^{l}_l} \). This is the case when collective solutions create local, rather than global, public goods whose boundaries and benefits are limited to group members as opposed to the entire population.

Fig. 4B indicates the group success rate as a function of loners’ externality (\( \rho \)), using actual choice data from our experiment for \( r^{\rho}_c \), \( r^{p}_d \), and, \( r^{l}_l \) in Eqs. 4 and 5. The figure illustrates several key results. When \( \rho = 1 \) (left endpoint), the group success rate with the individual option (blue) is lower than that without the individual option (red), corroborating with the previous argument that introducing an individual alternative to global public goods provisioning worsens the problem\(^6\). However, with a smaller \( \rho \) (toward the right endpoint), we can observe the relationship between the results with and without the individual option to reverse. This reaffirms the main claim of this study, namely, that voluntary participation aids collaboration in creating local public goods (Fig. 2B).

Note that the experimental data used to construct the graph were obtained under the no-externality scenario (\( \rho = 0 \)). We conjecture that variations in the parameter \( \rho \) are likely to affect not only the aggregation method (Eq. 4) but also the participants’ actions themselves. Therefore, the blue curve in the graph should be interpreted as an upper limit for the group success rate for each given \( \rho \). To summarize, the seemingly contradictory claims around the impact of an outside individual option on collaborative success can be integrated through the varying degrees of loners’ externality to the collective outcome, \( \rho \).

DISCUSSION

Collaboration enables us to achieve what no individual can accomplish alone. However, requiring multiple members’ efforts, it inevitably entails social uncertainty concerning cooperation and coordination\(^1\)\(^-\)\(^6\). Numerous studies have thus examined what behavioral tendencies, cognitive abilities, or bottom-up norms are needed to overcome these difficulties, as well as evolutionary pathways explaining why humans could have equipped themselves with such dispositions\(^7\)\(^-\)\(^17\),\(^42\),\(^43\). Yet, focusing attention on the free-riding problem, they have largely left out the mundane observation that collaborative efforts often take place among flexible group members who have gathered voluntarily.

Here, we have addressed this gap by modeling the collaboration as a voluntary threshold public goods game with an outside individual option. Using the game, we analyzed how voluntary participation can aid collaboration by promoting optimism among group members. We found that voluntary participation increased the likelihood of collaborative success via (1) the self-selection of optimistic cooperators into groups and (2) the formation of optimistic expectations among otherwise pessimistic defectors.

We note that theoretical and behavioral results consistent with the first self-selection mechanism have been reported in other settings\(^32\),\(^44\),\(^45\). For example, Orbell and colleagues\(^32\) had participants play a one-shot prisoner’s dilemma game under two conditions: the binary-choice condition, where the two players were obliged to choose between cooperation and defection, and the trinary-choice condition, with an additional exit option (more specifically, when one player selected the exit option, both players received a payoff lower than that of mutual cooperation but higher than that of mutual defection\(^32\)). Based on the observation that the population-level cooperation rate was greater in the latter condition, they argued that cooperators are less likely to exit than defectors. Although their use of a between-subjects design (rather than a within-subject design, as we used) seems to obscure this interpretation, their logic is similar to ours.

Most interestingly, we further observed that some pessimistic defectors formed more optimistic beliefs about others’ cooperativeness under voluntary participation, and opted for cooperation instead of becoming loners (Fig. 3C and D). Several previous studies using the repeated interaction paradigm have operationalized the exit option as the possibility to change partners and demonstrated that the combination of the exit option with some reputation mechanism can create an additional incentive for cooperation\(^36\),\(^52\). However, here, the increase in cooperation was driven by the intrinsic belief changes among participants, as opposed to the extrinsic behavioral-control system (i.e., partner selection; reputation mechanism) in those schemes.

Evaluating the outcome of the whole population, we found that voluntary participation significantly increased the group success rate (Fig. 2B) as well as efficiency (i.e., average payoff; Fig. 2C). As we have seen, the primary driver for players to defect within collaborative groups was fear rather than greed across the conditions. The results suggest that for successful collaboration, it is more critical to mitigate pessimistic defection than to prevent greedy defection, and voluntary participation just facilitates these dynamics.

Finally, in our view, whether outside individual options aid or hinder collaboration hinges on the degree of loners’
externality to the collective outcome. We argue that in most mundane instances under voluntary participation, group collaboration neither involves the entire population nor occurs within predetermined group boundaries. Consequently, the group consists solely of individuals who voluntarily opt in, and thus loners have little to no externalities to group endeavors.

There are several limitations about the scope of this paper. First, although this study focused on the idea that self-selection occurs on the basis of expectations about others’ cooperativeness, there may be other cognitive or motivational processes that influence self-selection in group endeavors. For example, if competence or confidence varies among individuals, different self-selection dynamics may operate, depending on specific incentive structures. In some cases, competent and/or confident members may be more likely to join the collaboration whereas they may pursue individual options if these outside options possibly yield relatively greater payoffs (see ref. for a review of the opt-in/opt-out mechanisms that operate in group collaborations). In this study, we assumed that the group collaboration, if successful, yields greater net profits than the individual option and that all members have equal competence/productivity affecting the success of the collaboration. In this sense, the situation addressed here should be considered a special case of voluntary participation. Nevertheless, we believe that our game incorporates the minimal incentive structures underlying collaborative situations and thus serves as a solid benchmark for further investigations.

Second, our study focuses on one-shot situations where decisions are made only once and does not explore situations where the decisions are repeated. Moreover, the mechanisms that promoted cooperation in the one-shot case relied on the variation of subjective expectations among players. This is not necessarily the case in repeated situations, as players observe their action histories in common and may adjust their expectations until they converge. Future research should investigate whether voluntary participation helps people maintain, not just initiate, cooperation in repeated interactions.

As an intensely social species, humans are frequently confronted with difficulties in and functionalities of collaboration. The current study suggests that the existence of an individual option renders participation in collaboration voluntary and can encourage individuals to pursue the collaborative endeavor with improved optimism.

**METHODS**

**Behavioral experiment**

ready graduated from the university and 14 participants who failed to participate in the experiment on time. Of the remaining 191 participants, 86 were male, 100 were female, 1 chose “other”, and the other 4 participants declined to answer. The mean age of participants was 22.8 years (SD = 2.7). For about 60 min of participation, participants were paid 2,124 JPY (SD = 316) on average (M = 18.47 USD, SD = 2.74). The experiment was approved by the Ethics Committee of the University of Tokyo.

We ran a total of six experimental sessions that lasted about an hour. In each session, around 30 (min: 30, max: 35) participants enrolled in the experiment from their own computers while being connected via Zoom. During the experiment, participants were kept anonymous and were not permitted to communicate with each other. At the beginning of each session, the experimenter read aloud the overall instructions to all participants while they viewed the instructions on their respective screens. We instructed participants that the monetary reward would be the sum of a constant completion fee (1,200 JPY) and a bonus based on their performance during the experiment. After giving informed consent, participants played the main task, proceeded to two additional tasks eliciting their risk and social preferences, and answered a postexperimental questionnaire.

For the main task, there were a total of six conditions, using a 2 (group participation: mandatory or voluntary) × 3 (threshold value: 2, 4, or 5) factorial within-subject design. All participants played each of the six conditions once. Participants were explicitly instructed that they were playing the game with around 30 other participants engaging in the experiment simultaneously. In each condition, subjects first read brief instructions about the rules of the respective game and took comprehension quizzes, during which they could ask the experimenter any questions via chat in Zoom. After answering all the comprehension quizzes correctly, they proceeded to the actual play. In the game, participants first estimated other participants’ actions (“How many of 30 other participants do you think will choose to cooperate, defect, or leave, respectively?”), indicated their confidence in their estimate on a scale of 0 to 100, and then decided on their own actions. There was no feedback about other participants’ decisions or resultant payoffs until the end of the whole experiment. The order of the six conditions was (partially) randomized across participants: Half of the participants played all three voluntary conditions first while the other half played all three mandatory conditions first, with the order of threshold values within the voluntary and mandatory conditions being randomized.

We incentivized estimations about others’ actions as well as participants’ own action selections. It was emphasized during the instruction that the bonus reward for the main task was set to increase as the participants estimated other participants’ actions more accurately and as they acquired more points from their own actions. Specifically, the bonus in the main task, vi, was determined randomly by either the participant’s estimation accuracy or the acquired points in one randomly selected condition.

If the estimation accuracy is selected:

**(Preprint)**
\[ v_i = 800 - \frac{80}{6} \times \sum_{x \in \mathbb{X}} \left( e_{i,x} - \frac{\sum_j I(x_j = x)}{\sum_j 1} \right), \]

and if the acquired point is selected: \[ v_i = 20 \times \pi_i \]

where \( e_{i,x} \) is the participant’s estimation about the number of other participants choosing the focal action \( x \), \( \sum_j I(x_j = x) \) corresponds to the actual number of participants (other than the participant themself) choosing the action, and \( \pi_i \) represents the points the participant earned in the focal condition. Note that for the estimation question, regardless of the exact number of actual participants other than the player themself (which ranged from 29 to 34), we asked participants to estimate actions of 30 others. The accuracy was then determined by comparing the ratio of each action to the total, as shown in the equation. The bonus was set to range from 0 to 800 JPY, regardless of whether the estimation accuracy or the acquired points was chosen.

After completing the six conditions in the main task, participants proceeded to two additional incentivized tasks designed to elicit their risk and social preferences, respectively. Specifically, they were asked to choose which lotteries to take and how to share a sum of money with another participant. From the participants’ answers to the lottery questions, we estimated their risk preference parameters assuming the constant relative risk aversion for their utility functions. From the answers to the sharing questions, we estimated their social preference using Fehr and Schmidt’s inequity aversion utility function. Both tasks were incentivized by telling participants that their monetary bonus would be determined by their decision on one of the randomly selected questions for each task (see the Supplementary Materials). At the end of the experiment, participants were paid the sum of the participation fee, the bonus from the main task, and the bonus from the additional task. See the Supplementary Materials for the actual protocols and the instruction slides used in the experiment.

**Preregistration**

The intended sample size, included variables, and main hypotheses were preregistered on Open Science Framework (https://osf.io/pdnwt) prior to the collection of any data.

**Statistical analyses**

To evaluate the variability of statistical values of interest (e.g., cooperation rate, group success rate, and efficiencies), we primarily employed bootstrap simulations. Since we used a within-subject design with order randomized both within and across sessions, our data are primarily clustered by participant (and not by session). To account for this, we used participant as the unit of resampling and calculate all the pertinent values (such as the cooperation rate and its difference between the voluntary and mandatory conditions) from that resampled data, instead of repeating independent resampling for each value. The exact calculation of the bootstrap 95% CIs proceeded as follows:

1. Resample \( N = 191 \) individuals from the experimental sample of \( N \) individuals with replacement.
2. Calculate the statistical values of interest (e.g., the difference in cooperation rates between the voluntary and mandatory conditions) from the resampled data.
3. Repeat Steps 1 and 2 for 1,000 iterations to obtain the distribution of the estimation of the statistics.
4. Report the range between the 2.5 and the 97.5 percentile of the empirical distribution as the bootstrap 95% CI of the focal statistics.

Error bars in the figures indicate the bootstrap 95% CI, unless otherwise stated. We interpret a 95% CI not containing 0 as evidence that a statistically significant difference exists.

Additionally, we introduced regression models in the following analyses. First, to assess how the cooperation (or defection) decisions in the mandatory conditions and leaving decisions in the voluntary conditions correlate with each other via subjective belief about others' cooperativeness, we evaluated three logistic regression models, each with a random intercept for participant: (1) defection in the mandatory conditions (i.e., defect = 1, cooperate = 0) as the dependent variable, and subjective belief in the mandatory conditions and threshold values as independent variables (i.e., fixed effects); (2) leaving in the voluntary conditions as the dependent variable, and subjective belief in the mandatory conditions and threshold values as independent variables; (3) leaving in the voluntary conditions as the dependent variable, and defection in the mandatory conditions (i.e., defect = 1, cooperate = 0) and threshold values as independent variables. Second, to explore economic or psychological traits (measured separately from the main task) that can partly account for inclinations toward the individual option rather than collaboration, we built a mixed-effect logistic regression with a random intercept of participants consisting of the leaving decision in the voluntary conditions as the dependent variable, and risk aversion, inequity aversion, interpersonal reactivity index, general trust, intolerance of uncertainty scale, the cognitive reflection test scores, and threshold values as independent variables. See the Supplementary Materials for detailed descriptions about each of these measures.

Last, we computed the ROC-AUC for each of the raw expectations about others' cooperation and the probability of their own cooperation being pivotal (which is calculated from raw expectation) in terms of the predictive power of action (see the Supplementary Materials for details about the ROC analysis). Notice that when the threshold is 5, the ROC-AUC of the raw expectation and the pivotal probability coincide. This is because the pivotal probability increases monotonically as the raw expectation increases (i.e., the order does not change) when the threshold is 5, and ROC-AUC is a rank metric that solely depends on the order of the predictions.
Therefore, we are only concerned with the threshold values of 2 and 4 in this comparison.

**ACKNOWLEDGMENTS**

We thank N. Mifune, N. Takahashi, and A. Ueshima for helpful comments on a previous draft, and A. Todd for editing the manuscript. We are also grateful for support from the Center for Experimental Social Sciences at Hokkaido University and the UTokyo Center for Integrative Science of Human Behavior.

**Funding:** This study was supported by the Japan Society for the Promotion of Science (grant no. JP16H06324 to T.K. and no. JP23JK0781 to R.M.), the Japan Science and Technology Agency CREST [grant no. JPMJCR17A4 (17941861)] to T.K., the Japanese Society of Social Psychology to R.M., and the Foundation for the Fusion of Science and Technology to R.M.

**Author contributions:**

Conceptualization: RM, NH, TK

Methodology: RM, NH, TK

Investigation: RM

Visualization: RM

Supervision: TK

Writing—original draft: RM, TK

Writing—review & editing: RM, NH, TK

**Competing interests:** The authors declare no competing interests.

**Data and materials availability:** The data and code to implement all analysis can be accessed at [https://github.com/ryutau/voluntary-collaboration](https://github.com/ryutau/voluntary-collaboration). All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

**References**


Department of Business Administration and Economics 393, (2017).

SUPPLEMENTARY MATERIALS

Supplementary Text

Difference in Nash equilibrium prediction between mandatory and voluntary participation

We first examine the threshold public goods game under mandatory participation (without the outside individual option). The game consists of five players who are each initially endowed with 10 points and can decide individually whether to contribute (C) their endowment (at a cost of 10 points) to the group or not (i.e., to defect, D; without the cost). If the number of players who contribute to the group reaches a certain threshold, \( q \) (here we consider only \( q = \{2, 4, 5\} \)), the public good is produced and all five players receive 30 extra points. However, if fewer than the threshold number of players contribute, the public good is not produced, and the contributions are not returned.

Since our situation does not allow players to coordinate with one another, here we consider only a symmetrical Nash equilibrium. Let us denote the probability of players' cooperation as \( p \). As in Eq. 1 in the main text, the expected payoff, \( \pi \), for each action (C or D) is computed as follows:

\[
\pi_C(p, q) = 30 \times \sum_{k=q-1}^{4} \binom{4}{k} p^k (1-p)^{4-k}, \\
\pi_D(p, q) = 10 + 30 \times \sum_{k=q}^{4} \binom{4}{k} p^k (1-p)^{4-k}
\] (S1)

Let us begin by searching for Nash equilibria consisting of pure strategies. The probability of players’ cooperation \( p^* \) is a symmetrical pure-strategy Nash equilibrium if and only if the conditions \( p^* = 0 \) and \( \pi_C(p^*, q) \leq \pi_D(p^*, q) \) or \( p^* = 1 \) and \( \pi_C(p^*, q) \geq \pi_D(p^*, q) \) hold. Solving these equations, \( p^* = 0 \) when \( q = 2, 4, 5 \) and \( p^* = 1 \) when \( q = 5 \) are the only pure-strategy equilibria.

Next, let us look for Nash equilibria that involve mixed strategies (\( 0 < p^* < 1 \)). Now \( p^* \) must satisfy \( \pi_C(p^*, q) - \pi_D(p^*, q) \equiv \Delta \pi = 0 \). Further, for the equilibrium to be stable, slight deviations from the equilibrium need to be pushed back. Namely, a slight increase (decrease) in \( p \) from \( p^* \) should result in a negative (positive) \( \Delta \pi \), that is, \( \frac{\partial \Delta \pi}{\partial p^*} < 0 \). Solving these, we obtain \( p^* \approx 0.41 \) for \( q = 2 \), and \( p^* \approx 0.88 \) for \( q = 4 \).

To sum up, under mandatory participation, the threshold public goods game with each threshold has one stable noncooperative equilibrium where no one contributes \( (p^* = 0) \) and one stable cooperative equilibrium where a certain proportion of the population does contribute \( (p^* > 0) \). Importantly, the expected payoff, \( E[\pi] \), for each equilibrium varies as follows:

<table>
<thead>
<tr>
<th>Threshold (( q ))</th>
<th>Noncooperative equilibrium</th>
<th>Cooperative equilibrium</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>( (p^*, E[\pi]) = (0, 10) )</td>
<td>( (p^*, E[\pi]) = (0.41, 26.3) )</td>
</tr>
<tr>
<td>4</td>
<td>( (p^*, E[\pi]) = (0, 10) )</td>
<td>( (p^*, E[\pi]) = (0.88, 27.9) )</td>
</tr>
<tr>
<td>5</td>
<td>( (p^*, E[\pi]) = (0, 10) )</td>
<td>( (p^*, E[\pi]) = (1, 30) )</td>
</tr>
</tbody>
</table>

Now, we examine how the introduction of the outside individual option under voluntary participation alters the prediction. Recall that the outside individual option secures a more certain but less lucrative payoff compared to collaboration: Players choosing to leave (L) earn 20 points in total (10 extra points + 10 points for the initial endowment) regardless of other players’ actions. Additionally, here we assume an infinite population and a nonzero probability of participation in groups (we do not consider a situation where there are too few players who opt in to groups and even a single group of five members cannot be formed).

Notice that across the three thresholds, the expected payoff under the noncooperative equilibrium is less than the 20 points secured for the individual option (left with the initial endowment of 10 points), whereas the expected payoff under the cooperative equilibrium is greater than 20 points. As a result, players expecting the noncooperative equilibrium within groups should opt out of groups to receive the greater payoff of 20 points from the individual option. In other words, voluntary participation in our setting eliminates the noncooperative equilibrium and thereby facilitates group collaboration toward cooperation.
Detailed Experimental Procedures and Instructions

Please refer to a separate file located at https://github.com/ryutau/voluntary-collaboration/blob/main/documents/experimental_instructions.pdf for the English translation of instruction slides that were presented to participants in the main task of the experiment.

Additional tasks eliciting participants’ economic and psychological characteristics

Apart from the main task (threshold public goods games), we included the following set of additional tasks and questionnaires to explore factors that possibly account for the individual heterogeneities in participants’ play in the main task:

1. Other-regarding preference\(^3\) (incentivized)
2. Risk preference\(^2\) (incentivized)
3. Interpersonal Reactivity Index\(^3,4\)
4. General trust\(^5\)
5. Cognitive Reflection Test\(^6,7\)
6. Intolerance of Uncertainty Scale\(^8\)

First, to measure participants’ other-regarding preference, we used the inequity aversion model developed by Fehr and Schmidt (FS model\(^1\)). The model presumes that a player’s utility is determined by the weighted average of their own payoff and the inequity of payoff between themselves and others, distinguishing between advantageous and disadvantageous inequity:

\[
 u_i = \pi_i - \alpha_i(\pi_j - \pi_i) - \beta_i(\pi_i - \pi_j)
\]

where \((\pi_i, \pi_j)\) denotes the payoff for the player and others, \(u_i\) is the resulting utility, and \(\alpha_i\) and \(\beta_i\) are individual-level parameters that quantify the player’s aversion to disadvantageous and advantageous inequity, respectively. As in He and Wu\(^9\), participants are repeatedly asked to choose their preferred allocations of piles of money between themselves and another participant chosen at random (independently of the main task). On the basis of their choices, we estimated \(\alpha\) and \(\beta\) for each participant.

Next, we presented participants with a gambling task to assess their risk attitudes: The participants were presented with a number of raffles with varying probabilities and prize amounts and were asked to choose one\(^2\). From their responses, we estimated participants’ risk attitudes assuming the constant relative risk aversion for their utility functions. Note that the Intolerance of Uncertainty Scale and the risk attitude measured from the gambling task differ not only methodologically (i.e., incentivized task vs. questionnaire) but also conceptually: The former is more concerned with attitudes toward mere possibility of a negative event irrespective of the probability of its occurrence than toward known probabilities\(^8\).

In both the allocation task and the gambling task, participants were directly incentivized to answer their preferences truthfully. Specifically, we randomly selected one target question from one target task and determined the participants’ actual bonus according to the exact procedure specified (e.g., dividing a pile of money with another randomly selected participant or holding the raffle they selected). Please refer to a separate file located at https://github.com/ryutau/voluntary-collaboration/blob/main/documents/experimental_instructions.pdf for the complete set of task items.

Comparison of raw expectations and pivotal probabilities in predicting cooperation within groups

In the threshold public goods game, players answered their own actions as well as their expectations about other players’ actions (“how likely are others to cooperate?”). How did their expectations determine their actions? Standard theories of expected utility maximization dictate that players should not respond to their raw expectations, but to the pivotal probability of their own decisions (i.e., both necessary and sufficient for the provision of collective benefits) calculated from them: Players with expecting higher pivotal probabilities are more likely to cooperate. However, we found that raw expectations were a better predictor of their cooperation: Players who expect other players to cooperate more are more likely to cooperate themselves.
Calculating the difference in the areas under the curves of the receiver operating characteristic [ROC-AUC] in predicting cooperation, they were significantly greater than zero across conditions ($\Delta AUC \equiv AUC_{\text{raw expectation}} - AUC_{\text{pivotal probability}}$): $q = 2$, mandatory: $\Delta AUC = 0.52$, CI $[0.39, 0.65]$; $q = 2$, voluntary: $\Delta AUC = 0.49$, CI $[0.30, 0.66]$; $q = 4$, mandatory: $\Delta AUC = 0.13$, CI $[0.07, 0.20]$; $q = 4$, voluntary: $\Delta AUC = 0.36$, CI $[-0.05, 0.63]$; when $q = 5$, the difference is 0 by definition). See also Fig. S4 for illustration.

**Exploratory analyses of correlations between participants individualistic choice and their economic and psychological characteristics**

We exploratorily analyzed participants’ economic and psychological characteristics (measured separately from the threshold public goods games) that may have been associated with the likelihood of going for the individual option in the threshold public goods games. As in previous research reporting little bearing of risk preference or distributive preference$^{10,11}$ (as measured by the social value orientation slider$^{12}$), our results also showed that neither risk aversion nor other-regarding preference (as measured via the inequality aversion utility model from Fehr and Schmidt$^1$) predicted participants’ choice of the individual option. Instead, participants with a lower tolerance for uncertainty$^8$ (as measured by questionnaires that include items such as “Unforeseen events upset me greatly”) were more likely to select the individual option under voluntary participation ($\beta=0.23$, CI $[0.004, 0.54]$). See Table S1 for the summary.
Supplementary Figure 1.

Voluntary cooperation induces higher cooperation rates via self-selection across a wide range of parameters of the Beta distribution. Resultant rates of cooperation (see Eqs. 2 and 3 in the main text) as a function of the distribution of beliefs, \( \phi(y) \). We assume that \( \phi(y) \sim \text{Beta}(a, b) \). The deeper orange indicates that the cooperation rate is closer to 1 and the lighter color shows the cooperation rate closer to 0. The panels correspond to the six conditions in the experiment. By comparing the results within each threshold value, we can confirm that voluntary participation induces higher cooperation rates compared to mandatory participation across a wide range of parameters.
Supplementary Figure 2.

**Breakdown of the participants choosing each action in each condition.** The pie chart depicts the number of participants who selected each action (orange: cooperate, blue: defect, and green: leave), with exact numbers in the wedges. The histogram shows a distribution of expectations about others’ cooperativeness ($\gamma$) colored by the action they chose.
Supplementary Figure 3.
Defection rates and noncooperation rates as a function of participants’ beliefs about others’ cooperativeness in the voluntary conditions. Regardless of the thresholds, individual defection rates and noncooperation (defection + leaving) rates decreased almost monotonically as participants’ expectations about others’ cooperativeness increased. Recall that the same pattern was also observed in the mandatory conditions (Fig. 3A in the main text). Note that when the threshold was 5, the defection rate (dashed lines) remains zero as there was no defector in this condition.
Supplementary Figure 4.

**Comparison of ROC curves of participants’ raw expectations and pivotal probabilities in predicting their cooperation decisions.** We compared goodness of each prediction by analyzing the respective receiver operating characteristic (ROC) curve used in the signal detection framework. The more the curve is situated above the diagonal, the better the prediction. Results indicate that the raw expectation predicts participants’ cooperation (orange curves) better than the pivotal probability (blue curves) regardless of the condition, with the exception of when the threshold is 5, where both values have the same order, resulting in identical ROC curves (thus two curves completely overlap, though displayed only in blue, in the upper-right panel). Observe that when the threshold is 2, the ROC curves drawn from the pivotal probability (blue curves) consistently lie below the diagonal, indicating the predictions were even inferior to completely random predictions. The statistical analysis in the supplementary text is based on the comparison of the area under the curve (AUC) of the two curves—the greater the AUC, the better the prediction. The lower right panel (when the threshold value is 5 and participation is voluntary) is empty because there were no defectors in this condition.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>2.5 percentile</th>
<th>97.5 percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.417</td>
<td>-1.007</td>
<td>0.174</td>
</tr>
<tr>
<td>Disadvantageous aversion ((\alpha) in FS model)</td>
<td>0.233</td>
<td>-0.012</td>
<td>0.676</td>
</tr>
<tr>
<td>Advantageous aversion ((\beta) in FS model)</td>
<td>-0.109</td>
<td>-0.297</td>
<td>0.018</td>
</tr>
<tr>
<td>Risk aversion</td>
<td>0.031</td>
<td>-0.174</td>
<td>0.256</td>
</tr>
<tr>
<td>Empathic concern (in the Interpersonal Reactivity Index (IRI; (41)))</td>
<td>-0.071</td>
<td>-0.329</td>
<td>0.186</td>
</tr>
<tr>
<td>Fantasy (in IRI)</td>
<td>0.106</td>
<td>-0.109</td>
<td>0.393</td>
</tr>
<tr>
<td>Personal distress (in IRI)</td>
<td>-0.13</td>
<td>-0.395</td>
<td>0.136</td>
</tr>
<tr>
<td>Perspective taking (in IRI)</td>
<td>0.202</td>
<td>-0.013</td>
<td>0.513</td>
</tr>
<tr>
<td><strong>Intolerance of uncertainty</strong></td>
<td>0.226</td>
<td><strong>0.004</strong></td>
<td>0.537</td>
</tr>
<tr>
<td>Cognitive reflection test score</td>
<td>0.062</td>
<td>-0.191</td>
<td>0.322</td>
</tr>
<tr>
<td>General trust</td>
<td>-0.191</td>
<td>-0.513</td>
<td>0.057</td>
</tr>
<tr>
<td>Age</td>
<td>-0.224</td>
<td><strong>-0.607</strong></td>
<td>-0.038</td>
</tr>
<tr>
<td>Gender (male)</td>
<td>0.239</td>
<td>-0.159</td>
<td>0.747</td>
</tr>
<tr>
<td>Gender (other)</td>
<td>-0.353</td>
<td>-2.154</td>
<td>0.866</td>
</tr>
<tr>
<td>Threshold</td>
<td>-0.052</td>
<td>-0.247</td>
<td>0.112</td>
</tr>
</tbody>
</table>

**Supplementary Table 1.**

**Estimated parameters of the mixed-effect logistic regression predicting the decisions to choose the individual option.** Independent variables include the threshold value and participants’ demographic data as well as economic, psychological, and cognitive characteristics that were measured separately from the main task. Red indicates items whose 95% confidence intervals do not contain 0. All numerical variables, except threshold values, are standardized across participants to have a mean of 0 and a standard deviation of 1; thus, the estimates shown are standardized coefficients.
Supplementary References