Strategic Behavior with Tight, Loose, and Polarized Norms

Eugen Dimant[1], Michele Gelfand[2], Anna Hochleitner[3], Silvia Sonderegger[4]

Date of latest draft: December 22, 2021

Abstract

A large body of literature has shown that social norms can be a powerful driver of human actions, and norm interventions often focus on shifting beliefs about average or majoritarian behavior. This work addresses a less studied aspect of norms, namely the role that their strength, tightness, and degree of polarization play. In the context of strategic decision-making, we investigate both theoretically and empirically how behavior responds to different distributions of one's peers' behavior. We focus on the difference between tight (i.e., characterized by low behavioral variance), loose (i.e., characterized by high behavioral variance) and polarized (i.e., characterized by U-shaped behavior) norm environments. Our results show that in addition to the mean, the variance, and shape of the observed behavior matters. In particular, we find that variance in observed behavior begets variance in one's own behavior. Tight, loose, and polarized norms are thus self-sustaining. In addition, we find that personal values matter more for behavior in loose and polarized norm environments.

1 Introduction

Empirical norms – the behavior adopted by other individuals – have been found to play a key role in shaping individual incentives in both strategic and non-strategic settings (Cialdini and Trost, 1998; Bicchieri, 2005, 2016; Bicchieri and Dimant, 2019; Dimant, 2019). Much of the existing literature in the social sciences focuses on the *average* or modal behavior when characterizing social norms (see e.g. Chaudhuri et al., 2006; Bicchieri and Xiao, 2009; Krupka and Weber, 2013; Feldhaus et al., 2019). In this paper, we argue that this approach misses out on an essential feature of norms—namely their *variance*. Consider for instance a collective action problem, and suppose that individuals in a given society contribute an average of 5 (on a scale from 0 to 10).

[1]University of Pennsylvania & CESifo, email: edimant@sas.upenn.edu
[2]Stanford University, email: gelfand1@stanford.edu
[3]University of Nottingham, email: anna.hochleitner@nottingham.ac.uk
[4]University of Nottingham, email: silvia.sonderegger@nottingham.ac.uk

This may reflect a situation where *everyone* contributes 5, a situation where each contribution level (0, 1, 2 .. 10) is selected by *one tenth* of the population, or a situation where *half* of the population contribute nothing and half contribute everything. Although all these scenarios generate an average contribution of 5, they clearly depict very different societies in terms of the underlying norm environment. Gelfand et al. (2011) speak in this context of *tight* versus *loose* norms and show that this distinction can help to understand systematic differences between cultures. Tight cultures are characterized by strong social norms and well-defined behavior, while loose cultures show a pattern of weak social norms and greater behavioral variance.

While there have been many studies of tight-loose in psychology, most are correlational (see e.g. Gelfand, 2021), and few studies have manipulated tight and loose norms and examined their impact on human cooperation in economic games (see Roos et al., 2015, for computational models of TL and the evolution of cooperation). In addition to the variance, we investigate another important feature of the distribution of behavior: its shape. More precisely, we are interested in polarized norms, i.e. situations where behavior is not single-peaked but follows a u-shaped distribution. Previous research suggests that this can have direct consequences on an individual's willingness to contribute to a public good (Henrich et al., 2001; Fehr and Fischbacher, 2003; Bowles and Gintis, 2013; Robbett and Matthews, 2021). Understanding how people react to polarized norms is thus an important issue, even more so in the face of increasing political polarization and the detrimental societal outcomes that are caused by it (Fiorina and Abrams, 2008; Iyengar and Westwood, 2015; Dimant, 2021; Gelfand et al., 2021).

To understand how the *shape* and *variance* of the distribution of one's peers' behaviors affects one's own behavior in strategic settings, we start with the premise that different distributions generate different degrees of strategic risk. In very tight environments, where behavior is tightly distributed and there is low variance, strategic risk is minimal, while in loose environments where there is high variance, the behavior of other individuals is less predictable, and strategic risk is substantial. We investigate how people respond to different levels of strategic risk. In our theoretical framework, we show that the answer depends on the exact specification of the utility function. People may react only to the mean of the distribution they face, ignoring strategic risk entirely (in the case of quadratic utility) or they may react to both the mean and the shape of the distribution (in the case of a linear-kinked utility). Which of these cases applies has important implications in terms of cultural sustainability. If best replies depend only on the mean of the distribution, a society will inevitably converge to the same distribution of behavior independently of whether the initial norm environment is tight or loose. Instead, in the case of linear-kinked utility, we show that optimal reactions will reproduce the initial distribution. When facing a tight norm environment, people will react in a tight way, while when facing a loose norm environment their reactions will be more spread out. Finally, when facing a polarized distribution, people will tend to react in a highly polarized way. This suggests that different norm environments (tight/loose/polarized) may be self-sustaining. Moreover, when faced with high levels of strategic uncertainty, we expect people to rely on their personal to guide their decisions. Hence, when an individual faces a loose environment with a lot of variance in behavior or faces a polarized environment, responses will be very different between people depending on their individual values (see Elster and Gelfand, 2021). Noteworthy, behavioral polarization occurs in spite of the fact that individual values are not polarized. In contrast, when faced with a tight environment, most individuals select similar contribution levels and personal values become a weaker predictor of behavior.

In a well-powered and pre-registered study, we utilize a representative sample of the U.S. population to test these different hypotheses empirically. We do so in the context of a variant of the established public goods game (PGG) with two players. Players receive a number of token at the beginning of the game and can decide to either keep them for themselves or to invest them in a public good that is then multiplied by a positive factor and shared equally among both players. This creates the classical dilemma between the socially optimal decision to allocate everything to the public good and the individually rational one to contribute zero (Ledyard, 1995; Fischbacher et al., 2001). To investigate the reaction of individuals to different norm-environments (tight/loose/polarized), we ensure that, before selecting their contribution, participants are informed about the distribution from which their co-player's contribution will be drawn. In a between-subject design we implement six different conditions that vary both the mean, dispersion and shape of the co-player's distribution. Our empirical results confirm the predictions under a linear-kinked utility function: the dispersion and shape of the co-player's distribution matters a lot for individual behavior. More precisely, we find that loose and tight norms are self-sustaining, meaning that if participants are faced with a higher dispersion, their contributions show a higher variance as well. Similarly, when participants are faced with polarized distributions, their reactions are highly divergent: they either choose to contribute a lot or very little. Our data also confirms that, in line with previous studies, being faced with a distribution with a higher mean induces higher contributions. This is an intuitive consequence of conditional cooperation and shows that individuals aim to match the contributions of others. Finally, in accordance with the theory, we find that personal values have a higher impact on individual contributions in loose/polarized environments relative to tight norm environments.

Our study contributes to the existing literature in various ways. Firstly, we provide causal evidence on how the tightness and looseness of a norm affects behavioral responses in a strategic context. Previous studies have either been correlational (see e.g. Gelfand, 2021) or focused on non-strategic environments (d'Adda et al., 2020). Our methodological approach allows us to decompose different mechanisms that explain why and how individuals respond to different shapes of norms. Secondly, we incorporate polarized norms in our framework by also looking at the *shape* of the distribution, which allows for a richer investigation of behavioral responses. Finally, we put the novel elicitation method by Dimant (2022) to a test, which allows us to measure not only beliefs about modal behavior or norms but also about its whole distribution in an incentive-compatible way. We thus put forward a more fine-grained measurement that helps to develop a better understanding of social norms and their impact on behavior. Our results also have practical implications for policy makers. For instance, the finding that personal values have a higher impact in loose and polarized environments suggests that behavioral change interventions should target them more in these circumstances. By contrast, if norms are tight, it may be more fruitful to focus on the behaviors of others.

The remainder of this paper is structured as follows. Section 2 provides an overview of related literature. In section 3, we outline our theoretical framework before section 4 describes the experimental design and states our hypotheses. Section 5 presents the empirical results, before section 6 concludes.

2 Related Literature

Our study contributes to a strand of research that explores the effect of norms on individual behavior. Many studies have shown in different contexts that providing information on what other people did in a given situation or what they think is appropriate influences individual decisions. This has been found both in non-strategic settings such as dictator games (Bicchieri and Xiao, 2009), voluntary payments (Shang and Croson, 2009; Feldhaus et al., 2019) or donations to charities (Dimant, 2019; Bicchieri et al., 2022), but also for strategic interactions such as PGGs (Chaudhuri et al., 2006; Kerr et al., 2009). The way most studies communicate the information about others is by focusing on the mean or modal behavior. Our study confirms that differences in means have a significant effect on subsequent decisions. However, we extend these findings by exploring differences between whole distributions. Furthermore, by investigating the normative dimension of the public goods game, we are building on seminal work by Bicchieri (2005) and Bicchieri and Chavez (2010) that proposes a theory of social norms based on empirical and normative expectations that together with personal values affect preferences and behavior. We find that the different distributions shift normative expectations in addition to beliefs about the co-players behavior. Moreover, we find that the importance of personal values depends on the relative tightness and looseness of the norm environment.

More recently, several studies have investigated norms in ambiguous environments. Fosgaard et al. (2020) for instance integrate random transfers in a standard dictator game and find that uncertainty about the source of transfers decreases aggregate norm compliance. Similarly, Ciranka and van den Bos (2020) develop and test a model in which social influence depends on individual uncertainty. Bicchieri et al. (2020) finally investigate both theoretically and experimentally environments in which signals about the applicable norms remain vague. The results suggest that individuals distort their interpretation of applicable norms to justify self-serving behavior. In our setup, however, we find no evidence of self-serving belief manipulation. Our work also has parallels to d'Adda et al. (2020). There, before selecting their action, dictators are shown different distributions of normative views taken from a previous study (baseline, low mean and high variance). Their results show that when the variance of the shown distribution is lower, the variance of dictator contributions decreases as well. While building on their findings, our study differs from their set-up in several important ways. Firstly, the environment we analyze is one where individuals interact strategically and where each participant is confronted with strategic uncertainty. The psychological mechanism behind our results is therefore fundamentally different. Moreover, we consider a wider range of distributions and vary not only their variance but also their shape. We believe that the case of polarized norms is a particularly interesting one to study in the face of an increasing discussion about social polarization (Dimant, 2021). Finally, by measuring behavior both before and after participants are shown the distributions, we are able to identify within-participant shifts and control for baseline behavior in our analysis.

By investigating how polarized norms affect individuals' decisions and beliefs this work is moreover linked to the literature on polarization and its effects on behavior and preferences (see e.g. Fiorina and Abrams, 2008; Iyengar and Westwood, 2015; McConnell et al., 2018; Iyengar et al., 2019; Bursztyn et al., 2020; Dimant, 2021; Robbett and Matthews, 2021). Bénabou and Tirole (2016) argue that polarization is maintained as people interpret and process the same information in very different ways along with a growing divergence in beliefs. We show experimentally, that observing polarized norms in fact leads to very heterogeneous reactions with different people focusing on different parts of the distribution. This translates into both polarized beliefs and actions and thus a re-enforcement of the initial norm.

Finally, as we explore how different information about the behavior of others affects contributions in a PGG, we add to a literature that stresses the importance of reciprocity and conditional cooperation when trying to understand contributions to a public good (see e.g. Fischbacher and Gächter, 2010; Bowles and Gintis, 2013; Gächter et al., 2017).

3 Theoretical framework

Let x_i denote own contribution, x_j is co-player's contribution and X is the endowment. It is well known that, in strategic environments, reciprocity plays an important role in determining an individual's choice of action. The presence of reciprocity motives has also been extensively documented by the literature on public good games (see e.g. Fischbacher and Gächter, 2010; Bowles and Gintis, 2013; Gächter et al., 2017). The underlying idea is that individuals incur a psychological cost whenever their contribution differs from that of their co-player. Consequently, they adapt their behavior to the behavior of their co-player. We are interested in a setup where individuals do not observe their co-player's behavior before selecting their action, but know that the action of their co-player is drawn from a distribution $f(x_j)$ on $[0, \overline{x}]$ with mean μ_x and variance σ_x^2. Intuitively, this depicts the idea that, when engaging in everyday interactions, people may not know what their counterpart will do, but may be aware of the typical distribution of behavior within society. Since their co-player's contribution is unobserved at the time when they choose their action, agents are exposed to *strategic risk*. Their choice needs to trade off the risk of contributing too little (relative to their co-player) and the risk of contributing too much. These competing motives will determine the optimal contribution for an individual when confronted with a distribution of co-player's actions.

How do people react to environments with loose vs tight norms? How do they react to polarized norms? These are important questions that help shed light on what we may expect to observe within society. Suppose for instance that, when confronted with polarized norms – i.e., the distribution from which their co-player action is drawn is U-shaped – individuals tend to react by selecting an action that is middle of the road, in order to hedge the amount of strategic risk they are exposed to. This would suggests that societies characterized by polarized norms will tend to unravel, since the optimal reaction by individuals when confronted with polarized norms is to choose actions that are tightly distributed rather than polarized. If by contrast individuals show very heterogeneous reactions to polarized norms, the existing equilibrium is self-sustaining.

The optimal reaction to a given distribution of co-player's actions depends on the nature of the utility function. There are two natural ways in which the agents' concern for matching their co-player's action can be captured formally: (i) through a quadratic loss function – where the psychological cost incurred is proportional to the square of the difference between the two contributions, or (ii) through a linear-kinked loss function, where the cost is proportional to the absolute value of the difference between the two contributions. Quadratic loss functions are commonly used for instance in models of conformity or coordination (see e.g. Kandel and Lazear, 1992). The linear-kinked loss function was introduced by Fehr and Schmidt (1999) and has been widely used ever since.

3.1 Quadratic loss function

Consider the following stylized model of reciprocal preferences:

$$u_i = X - x_i + \gamma(x_i + x_j) - \frac{\eta_i}{2}(x_i - x_j)^2 \tag{1}$$

where $X - x_i + \gamma(x_i + x_j)$ is material payoff, $\frac{1}{2} < \gamma < 1$ and η_i parametrizes i's reciprocity concerns.[1] The last term in (1) captures the desire to minimize the psychological costs incurred whenever the player's contribution differs from that of the co-player. Each individual i selects x_i to maximize their expected utility, where the expectation is taken with respect to x_j. We denote i's optimal contribution as x_i^*.

Lemma 1: *When utility is given by (1), we have (i) $x_i^* = 0$ if $\eta_i < (1-\gamma)/\mu_x$, (ii) $x_i^* = \mu_x - (1-\gamma)\frac{1}{\eta_i}$ otherwise.*

In other words: i's contribution (when positive) is equal to j's expected contribution minus a constant which is decreasing in i's concern for reciprocity. The implication is that,

Proposition 1: *When utility is given by (1), the optimal contribution depends on the distribution of co-player's contribution only through the distribution's mean μ_x.*

Intuitively, the quadratic loss function implies that individuals are *averse to strategic risk.* They therefore choose their contribution to minimize the strategic risk they are exposed to. The optimal solution to this problem indexes i's contribution to μ_x, the co-player's mean contribution. This ensures that the difference between x_i and x_j is never too large. Crucially, it implies that i's choice *only* depends on $f(x_j)$ through μ_x, and is independent of the other features of the distribution of co-player's behavior.

3.2 Linear-kinked loss function

Suppose now that utility is

$$u_i = X - x_i + \gamma(x_i + x_j) - \alpha_i (x_i - x_j)\,|_{x_j < x_i} - \beta_i (x_i - x_j)\,|_{x_j > x_i} \tag{2}$$

[1] Note that, although η_i will typically depend on the degree of intentionality in the co-player's action, for the purpose of our design, where action intentionality is the same across all treatments, it can be treated as constant.

This utility function differs from (1) in that the psychological loss incurred by individuals is proportional to the *absolute value* of the difference between their contribution and that of their co-player. The parameter α_i (resp., β_i) measures the marginal disutility obtained from selecting a contribution that exceeds (resp., is lower than) the co-player's contribution.

Lemma 2: *Let* $\varphi_i \equiv \frac{\beta_i-(1-\gamma)}{\alpha_i+\beta_i}$. *When utility is given by (2), we have (i)* $x_i^* = 0$ *if* $\varphi_i \leq 0$, *(ii)* $x_i^* = \overline{x}$ *if* $\varphi_i \geq 1_i$, *(iii)* x_i^* *satisfies* $F(x_i^*) = \varphi_i$ *otherwise.*

Figure 1 represents the function $F(x)$ for the case of (i) single-peaked distributions and a (ii) polarized (U-shaped) distribution. In panel (i), the solid line represents a distribution with a smaller variance compared to the dashed line. The horizontal straight lines represent φ_i.

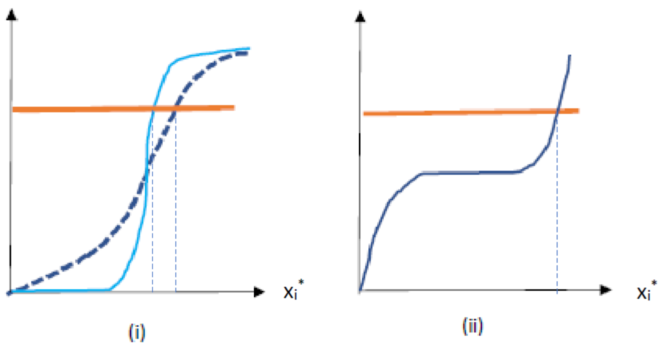


Figure 1: F(x) for single peaked (i) and polarized (ii) distributions

As can be seen from the figure, the point where $F(x)$ and φ_i cross depends on the nature of the distribution of co-player contributions. For instance, when $f(.)$ is polarized, $F(x)$ is steep at the extremes and flat in the middle. This implies that, typically, $F(x)$ and φ_i will cross when x takes extreme values – either very low or very high (panel (ii) of the figure illustrates the latter possibility). When facing a polarized distribution, individuals thus exhibit strategic-risk taking behavior: they prefer to take a gamble and risk ending up in a completely mismatched position vis-à-vis their co-player rather than opting for a "middle of the road" contribution level which would minimize risk.

In contrast, when the distribution of co-player contributions is single-peaked, $F(x)$ is flat at the extremes and steep in the middle. Consequently, $F(x)$ and φ_i will tend to cross when x takes intermediate values. As the variance of $f(.)$ increases, though, x_i^* will tend to become progressively more extreme, as can be seen by comparing the solid and the dashed lines in panel (i) of Figure 1. These observations lead to,

Proposition 2: *When utility is given by (2), the optimal contribution will typically depend on the variance and shape of the distribution of co-player's contribution in addition to the distribution's mean μ_x.*

The following result formalizes the notion that, as the variance of co-player contribution increases, individuals tend to select more extreme contributions. Consider two distributions f_0 and f_1 with the same mean μ_x and suppose that f_0 is single-crossing stochastic dominant over f_1 (Machina and Pratt, 1997). Note that this implies that f_1 is a mean-preserving spread of f_0. Denoting the optimal contribution under f_k as x^*_{ik}, the following holds.

Proposition 2a: *There exists $z \in (0, \overline{x})$such that the following holds; (i) if $x^*_{i0} \in (0, z)$: $x^*_{i1} < x^*_{i0}$; (ii) if $x^*_{i0} \in (z, \overline{x}) : x^*_{i1} > x^*_{i0}$.*

Finally, from Lemma 2 is is straightforward to see that,

Proposition 2b: *Consider f_2 and f_3 such that $F_2(x) < F_3(x) \, \forall x$. Then, $x^*_{i2} \leq x^*_{i3}$ with strict inequality whenever at least one of x^*_{i2} and x^*_{i3} is interior.*

This corresponds to the case where the distributions of co-player contributions have different means (but the same variance).

3.3 The role of personal values

We now allow agents to also be concerned with abiding to their own personal values when selecting their contribution, in addition to their desire to match their co-player's action. Let x^a_i represents i's perception of what constitutes "the right thing to do" (the "appropriate action"), and suppose that individuals suffer a psychological loss when their contribution differs from x^a_i. Concern for own personal values introduces two additional channels through which the distribution of co-player's actions may influence the agent's contribution. First, they may affect the agent's perception of what is appropriate. For instance, the distribution of co-player contributions might reveal information about what others consider to be the right thing to do, and this may affect what i sees as appropriate. Second, in the case of a linear-kinked loss function, the shape of the distribution of co-player's behavior may mediate the extent to which the appropriate action x^a_i affects the individual's optimal contribution. To see this, let us augment the linear-kinked utility described in (2) with appropriateness concerns as follows.

$$u_i = X - x_i + \gamma(x_i + x_j) - \alpha_i (x_i - x_j) \mid_{x_j < x_i} - \beta_i (x_i - x_j) \mid_{x_j > x_i} - \frac{\delta_i}{2} (x_i - x_i^a)^2 \qquad (3)$$

where δ_i parametrizes the importance that i ascribes to doing the right thing.[2]

Lemma 3: Let $\phi_i \equiv \beta_i + \delta_i x_i^a - (1 - \gamma)$. *When utility is given by (3), we have (i)* $x_i^* = 0$ *if* $\phi_i \leq 0$, *(ii)* $x_i^* = \overline{x}$ *if* $\phi_i \geq \delta_i \overline{x} + \alpha_i + \beta_i$, *(iii)* x_i^* *given by* $\delta_i x_i^* + F(x_i^*)(\alpha_i + \beta_i) = \phi_i$ *otherwise.*

Proposition 3: *Suppose that utility is given by (3). Whenever* x_i^**is interior, we have*

$$\frac{\partial x_i^*}{\partial x_i^a} = \frac{\delta_i}{\delta_i + f(x_i^*)(\alpha_i + \beta_i)}$$

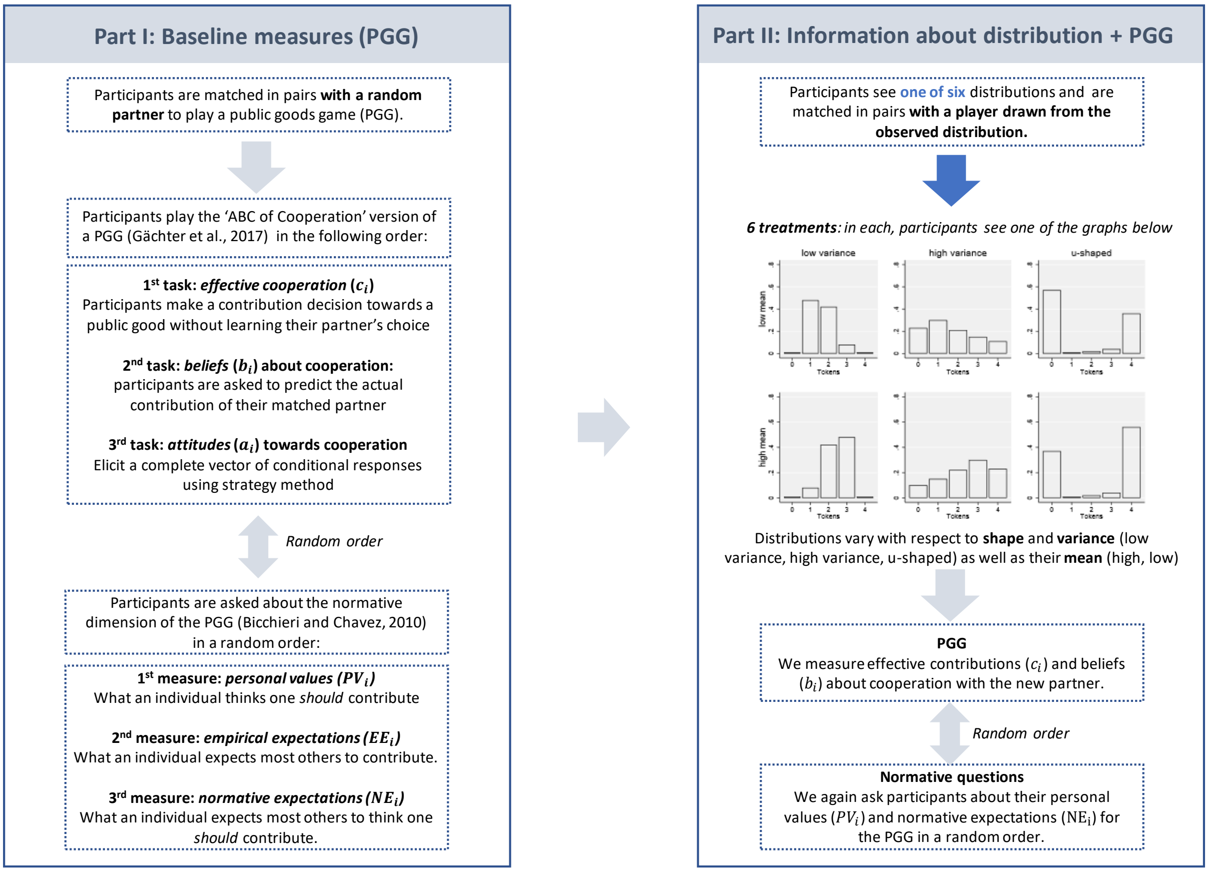
In other words, the optimal contribution is more responsive to differences in an agent's personal values when the $F(.)$ function is flatter – reflecting higher variance in the distribution of co-player's actions.

4 Experimental Design

To empirically test the propositions derived in our theoretical framework, we design an experiment that exogeneously varies the mean and shape of the co-player's behavior in a PGG. More concretely, the experiment consists of two parts. Participants learn the details of the second part only upon completion of the first. Part I elicits baseline behavior and (normative) beliefs for the PGG. At the beginning of Part II, participants observe the distribution of behavior from a previous session and are informed that their co-player for part II will be randomly drawn from this distribution. We then ask them to again make a choice for the PGG and measure their beliefs. The distributions vary along two dimensions: *mean* (high, low) and *variance* (high, low, u-shaped). Figure 2 illustrates the design. Tying this design decision back to our motivation, the low variance conditions are an example of a tight norm, while high variance conditions are an example of loose norms. The u-shaped conditions, finally, depict the case of a polarized norm.

[2] The results described to the previous subsections are robust to the inclusion of concern for personal values. Details available upon request.

Figure 2: Overview of the experimental design



4.1 Hypotheses

Before describing the experimental design in more detail, we lay out how our theoretical framework informs the behavior we expect to see in the experiment. The experiment and our hypotheses were pre-registered. There are three main hypotheses that we want to test. Firstly, independent of the variance and the shape of the distribution we expect contributions to be higher in the high mean conditions. This is in line with the prevalence of conditional cooperation in PGGs and shows a desire to match the contributions of the other player.

Hypothesis 1: *Individual contributions depend on the distribution's mean, implying that contributions in high mean conditions are higher than in low mean conditions.*

The effect of tight, loose, and polarized norms depends on the assumptions about the underlying loss function. As *proposition 1* indicates, under a quadratic loss function we expect participants to only react to changes in the mean of the observed distribution. If we assume a linear-kinked loss function, by contrast, we expect in line with *proposition 2* that the shape of the distribution matters, too. Our theoretical framework also predicts the direction of the effect. We expect higher dispersion to lead to more extreme contributions and thus a larger behavioral variance. If the observed norm is polarized, we expect to see a polarization of contributions.

Hypothesis 2a: *Assuming a quadratic loss function, participants only react to the mean of the distribution. There is no effect of loose, polarized or tight norms on contribution behavior.*

Hypothesis 2b: *Assuming a linear-kinked loss function, loose norms lead to a larger variance in contribution behavior. Moreover, when loose norms are polarized (U-shaped) they generate polarized behavior.*

Finally, *proposition 3* shows that if we take into account personal values or what individuals see as the appropriate action, the shape of the distribution mediates the extent to which personal values affect behavior. Under the assumption of a linear-kinked loss function, we expect a higher influence of personal values on behavior under loose/polarized than under tight norms.

Hypothesis 3: *Assuming a linear-kinked loss function, personal values have a higher impact on contribution decisions in loose and polarized compared to tight norm environments.*

4.2 Treatment conditions

Our experiment is based on a between-subject design, where participants are randomly assigned to one of six treatment conditions. More precisely, we use a 2x3 design, resulting in six experimental conditions that combine high and low means with three different distribution shapes (low variance, high variance, u-shaped). Table 1 gives an overview of the different treatment conditions, including their mean (μ) and variance (σ^2).

	Single-peaked				*Double-peaked*	
	Low variance		**High variance**		**U-shaped**	
	mean	var	mean	var	mean	var
High mean	1.6	0.5	1.6	1.6	1.6	3.6
Low mean	2.4	0.5	2.4	1.6	2.4	3.6

Table 1: Experimental conditions

While part I is identical across treatments, participants observe different distributions at the beginning of part II.[3] The latter are constructed through non-random sampling of previous session.[4] Participants are told in the experiment that we invited several hundred people before them to play the same game and that we used their answers to construct six different sub-groups. Moreover, they know that in the second part of the experiment, we will randomly show them the behavior of one of these sub-groups. We are thus completely transparent with participants and acknowledge that the distributions do not represent overall behavior in a PGG, but only represent the behavior of our constructed sub-groups. This allows us to use non-random sampling without being deceptive. Most importantly, participants know that their co-player's contribution choice will be drawn randomly from the observed distribution and thus affect their payoffs. Likewise, their decisions affect the payoffs of the randomly drawn player. We illustrate the distributions of all six conditions that participants saw at random in Figure 3.[5]

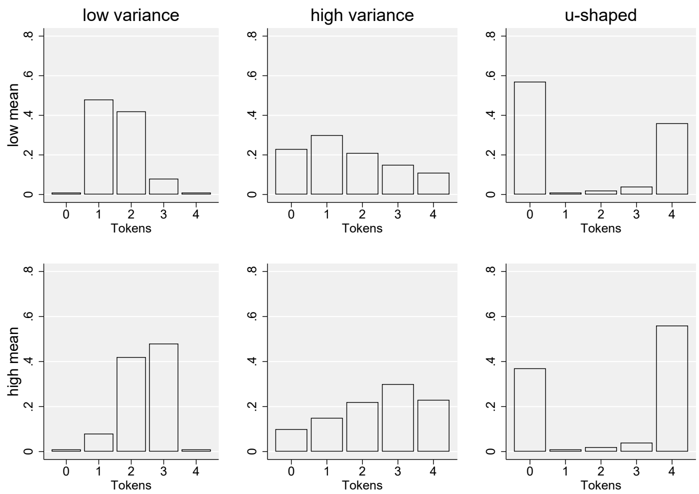


Figure 3: Experimental conditions: distribution of co-player's contribution

4.3 Measuring contributions and beliefs

In Part I, we elicit baseline measures for individual behavior as well as perceived norms in a standard PGG (Ledyard, 1995; Fischbacher et al., 2001). This allows us to later control

[3]See appendix B for experimental instructions.
[4]A similar approach is used for example by Frey and Meier (2004), Bicchieri and Xiao (2009), Krupka and Weber (2009), and Bursztyn et al. (2020).
[5]For an exact representation of how participants receive the information about distributions see appendix B.

for baseline behavior in our analysis. We use a two-player variant of the PGG in which each participant can contribute up to four token to the public good. Any token invested in the public good is then multiplied by 1.4 and shared between both participants. As mentioned above, the game embodies the classic tension between private and collective interest: while fully contributing to the public good maximizes joint payoffs, each player's self-interest is maximized by contributing nothing. To disentangle the underlying motives to contribute, we follow the 'ABC of cooperation' (**A**ttitudes-**B**eliefs-**C**ontribution) method developped by Gächter et al. (2017). The ABC method embodies three distinct elicitations: a one-shot sequential PGG played with the strategy method to measure attitudes of cooperation (A), a belief-elicitation task to measure expectations of others' contribution (B), and a one-shot simultaneous PGG played with the direct response method to measure effective contributions (C). We always first measure contributions, followed by beliefs and attitudes. The elicited attitudes give us a conditional contribution vector that can then be used to classify participants into different *cooperation types*. Following the seminal paper by Fischbacher et al. (2001) and the refinement by Thöni and Volk (2018) we distinguish the following five types:

1. *Free riders*: contribute 0 token, independent of the other's contribution.
2. *Conditional cooperators*: show either a monotonically increasing pattern or a strong positive correlation between own and other's contribution (Pearson $\rho \geq 0.5$).
3. *Unconditional cooperators*: contribute the same positive amount, independent of the other's contribution.
4. *Triangle cooperators*: reach a maximum contribution at a middle value x. Contributions either show a strong positive correlation to the left and a strong negative correlation to the right or are monotonically increasing until x, then decreasing.
5. *Others*: can not be classified using the criteria specified above.

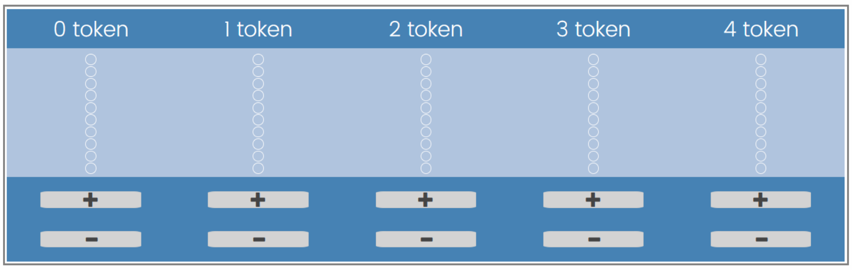
In addition to contributions, we are also interested in the perceived normative dimension of the PGG. Following Bicchieri and Chavez (2010), we measure both personal values (PV), i.e. what an individual thinks one *should* do[6], as well as empirical and normative expectations (EE and

[6]The literature also refers to this as personal normative beliefs.

NE) about what an individual expects most others to do and most others to think one *should* do. All three measures (PV, EE, NE) are elicited in a randomized order.[7] In part II of the experiment, we again ask for individual contributions and beliefs, as well as participants' PVs and EEs. The only difference to part I is that we are not using the strategy method and that the co-player is not chosen randomly but drawn from the shown distribution. Contribution decisions in both part I and II are incentivized and we randomize whether participants first make their contribution decisions or indicate their normative views.

As we are interested in the effect of the variance of norms, we do not only capture beliefs about modal behavior, but about the whole distribution of the participants' beliefs in an incentive compatible way. To do so, we follow an approach introduced by Dimant (2022) and ask participants to allocate points across all possible outcomes (see Figure 4). The more likely participants think an option is, the more points they should allocate to it. To ensure incentive compatibility payoffs are calculated using a quadratic scoring rule (QSR): $Q_j(p) = \alpha + 2\beta * p_j - \beta * \sum_{i=1}^{n}(p_i)^2$, where p_j is the probability a participant assigns to the true option.[8] We use this approach for both the beliefs about the co-player's contribution as well as the normative and empirical expectations.[9] To avoid hedging between the different questions, we randomly select one of them to be payoff relevant at the end of the experiment.

Figure 4: Belief elicitation screen



Note: Participants have to allocate a total of 10 points across all available options. The more likely they think an option is the more points they should allocate to it.

[7]When eliciting PV and EE, we ask what an individual considers to be the most *appropriate* action or the action that most people would agree upon as being "correct" or "moral" (see Bašić and Verrina, 2020).
[8]In the experiment we set $\alpha = \beta = 0.5$, which implies that participants earn between \$0 and \$1 depending on their stated beliefs and the truth.
[9]In order to make the rule as easy to understand as possible, we adapt the instructions developed by Artinger et al. (2010) that allow a transparent representation of the QSR even for non-binary decisions and use an intuitive interface to make decisions (see Quentin, 2016). For details see instructions in appendix B.

4.4 Sample and data collection

We programmed the experiment using Qualtrics (2005) and recruited participants online via Prolific Academic in December 2021. In total, we recruited a sample of about 1200 US participants that is representative in terms of age, gender and ethnicity (see Table 2 for observations per treatment). The chosen sample size was determined using data from a pilot and allows us to detect treatment effects at a 5% significance level with 90% power. On average participants needed 17 minutes to complete the study and earned $3.20.

	Single-peaked		**U-shaped**
	low variance	high variance	u-shaped
high mean	201	201	201
low mean	199	200	201

Table 2: Sample size and experimental conditions

Moreover, to construct the six distributions, we collected data from 685 MTurkers in September 2021. They initially received a show-up fee and then earned an additional bonus depending on the decisions of the participants in the main experiment.

After completing the main experiment, participants fill out an ex-post survey that provides further controls for our analysis. In terms of demographics and general characteristics we collect information on participant's gender, age, highest education level, trust, attitudes towards risk (Dohmen et al., 2011), as well as negative and positive reciprocity (Falk et al., 2016). In addition, we ask questions about the observed distributions and the experiment. Here we ask participants how they perceive average contributions as well as their variance in the observed distribution (low, medium, high) and how difficult they found interpreting the graph (on a scale from 1 to 7). As we tell participants that they see observations from one of the six sub-groups we constructed from previous sessions, we also ask how representative they think this behavior is on a scale from 1 to 7. Finally, to proxy their aversion of contributing too much/ too little, we ask them on a scale from 1-7 "How upset would you be if you invested everything in the group account and discovered that the participant you have been matched with invested nothing?" (*sucker aversion*) and "How ashamed would you be if you invested nothing in the group account and discovered that the participant you have been matched with invested everything?" (*free-rider aversion*).

5 Results

5.1 Baseline measures

We first provide an overview of our baseline measures, both with respect to the ABC method as well as personal values and perceived norms (EEs and NEs).

The conditional contribution schedule reveals a strong pattern of conditional cooperation among participants, which is in line with existing literature. Following the type definitions developped by Fischbacher et al. (2001) and Thöni and Volk (2018), we classify 84% of all participants as conditional contributors.[10] On average, participants contributed 2.6 out of their 4 token to the public good. Beliefs about the co-player's contribution were on average slightly more pessimistic (2.3 token). Moreover, there was quite a high variation in responses, stressing the strategic uncertainty of the PGG. As the left graph in Figure 5 shows, the most common beliefs were thereby that the other contributes 2 or 4 token.

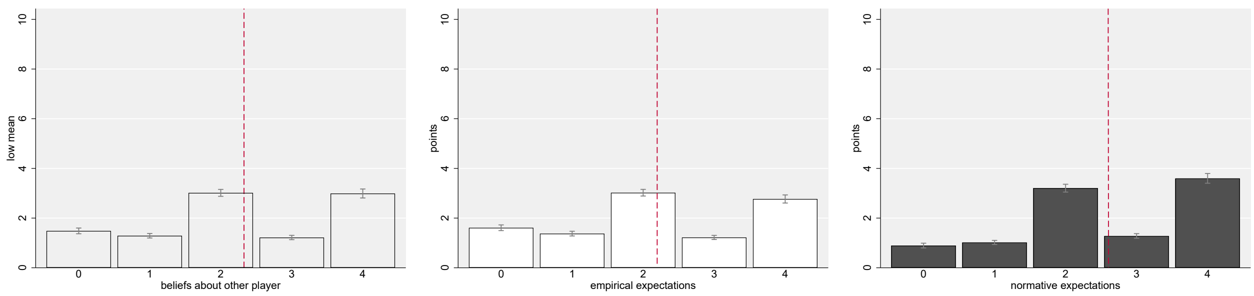


Figure 5: Beliefs in Part I

When it comes to empirical and normative expectations, we see a very similar pattern (see Figure 5). Overall, our results show that participants expect a wide range of behavior and have no precise NE or EE, indicating that ex-ante there is no clearly defined norm in the PGG. Not surprisingly, when asked about what most other people actually contribute participants answer in the same way as when asked about their co-player's likely contribution. NEs, are shifted slightly to the right, indicating that people think other's contribute less than they say one should. The same also holds true for the participants themselves. On average participants state that one should

[10]The rest consists of 5% unconditional contributors, 3% free-riders, 5% triangle contributors and 4% others.

contribute 2.9 token to the public good. This is significantly higher than the actual contributions we observe in part I.

5.2 Hypothesis 1: Effect of high and low means

We first examine whether our treatment manipulations worked. In the ex-post survey we asked participants about their perceptions of the shown distribution. As Figure 6a shows, participants in high mean conditions also have a significantly higher perception of the mean ($p < 0.01$).[11] In terms of the variance (Figure 6b), participants perceive the high variance and the u-shaped condition as significantly more varying than the low variance conditions ($p < 0.01$).

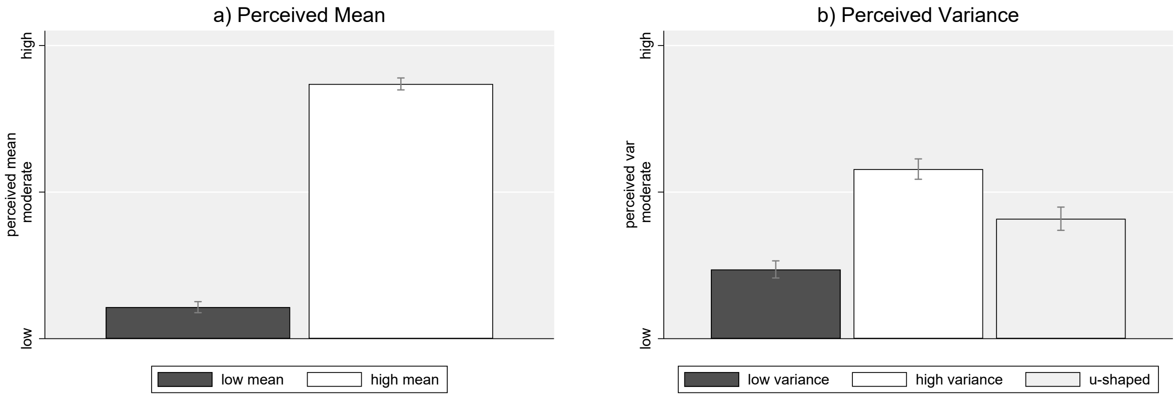


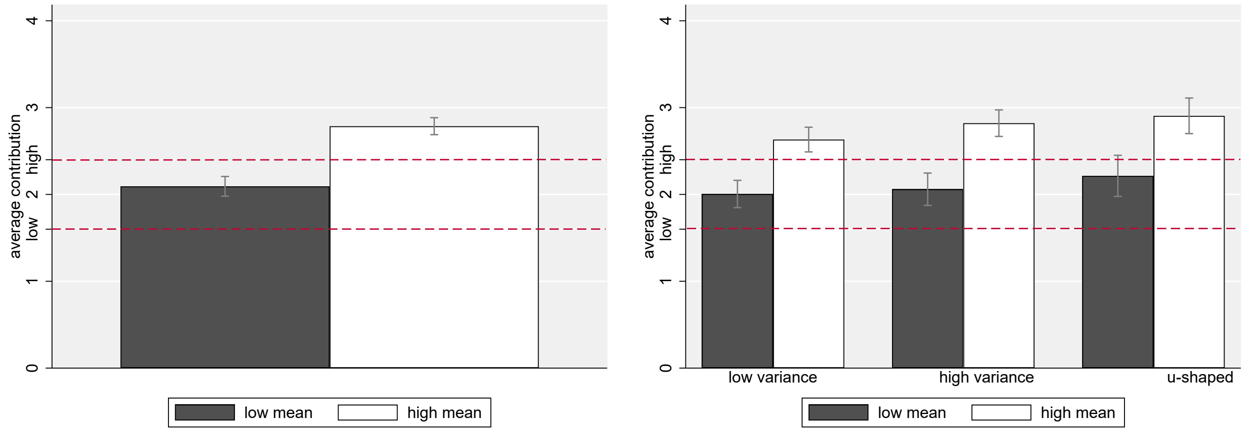
Figure 6: Manipulation checks

As noted above, there is an absence of a clearly defined norm and a strong pattern of conditional cooperation in Part I. These observations suggest that the distribution of the co-player's behavior should play an important role in determining individual contributions. When first looking at the difference between high and low mean conditions, we find that in line with previous literature contributions are significantly higher after observing high mean distributions ($p < 0.01$). This is both true when pooling the data and when separately testing the effect for each shape and variance (see Figure 7). Moreover, we find that the mean has a similar effect for all variances, as the diff-in-diff between bars is not statistically significant.[12]

[11] If not stated otherwise p-values refer to a two-sided t-test.

[12] Table A.1 in appendix A shows the results of regressing contributions in part II on treatment indicators (mean and variance) and their interactions.

Figure 7: Effect of high or low mean on contributions (left: overall, right: by variance)



Note: The dashed lines show the mean of the observed distributions (high = 2.4, low = 1.6)

To test **hypotheses 1** formally and control for individual differences we estimate:

$$C_i = \beta_0 + \beta_1 Mean_i + \beta_2 Variance_i (+\beta_3 X_i) + \epsilon_i \tag{4}$$

where C_i is the number of token contributed by individual i, $Mean_i$ is a binary variable that takes the value 0 if participants are assigned to the low mean condition and 1 if they are assigned to the high mean condition. Var_i is a categorical variable that is 0 if participants are assigned to the low variance condition, 1 if they are assigned to the high variance condition, and 2 if they are assigned to the u-shaped condition. X_i is a vector of individual controls. As contributions in our experiment are limited between 0 and 4 token, we also run tobit regressions as a robustness check. As Table 3 shows, our finding that contributions are significantly higher in the high mean conditions holds across all specifications, even after controlling for decisions in part I.

Result 1: *Participants react to the mean of the distribution. When confronted with higher average contribution levels, they contribute significantly more.*

One reason why different distributions cause different average contributions is that they successfully shift beliefs about the co-player's behavior. Figure 8 shows that when comparing the cumulative distributions of average beliefs in high and low mean conditions the latter are lower throughout.[13]

[13] A Kolmogorov-Smirnov test confirms that the distributions are statistically different at the 1% level.

Table 3: Effect of high and low mean conditions. Dependent variable = contribution in part II

	OLS			Tobit		
	(1)	(2)	(3)	(4)	(5)	(6)
high mean	0.69***	0.70***	0.66***	0.99***	0.99***	0.93***
	(0.08)	(0.06)	(0.06)	(0.14)	(0.11)	(0.10)
Variance (baseline = low)						
high variance	0.12	0.14**	0.16**	0.22	0.25**	0.26**
	(0.08)	(0.07)	(0.07)	(0.16)	(0.13)	(0.12)
u-shaped	0.24**	0.29***	0.27***	0.61***	0.66***	0.61***
	(0.10)	(0.08)	(0.08)	(0.17)	(0.13)	(0.13)
Baseline controls						
contribution in part I		0.49***	0.45***		0.80***	0.73***
		(0.04)	(0.04)		(0.06)	(0.06)
belief in part I		0.19***	0.16***		0.35***	0.30***
		(0.05)	(0.05)		(0.08)	(0.08)
personal values in part I		0.04	0.03		0.06	0.04
		(0.04)	(0.04)		(0.06)	(0.06)
Constant	1.97***	0.15	-0.06	1.98***	-1.09***	-1.75***
	(0.07)	(0.10)	(0.39)	(0.13)	(0.20)	(0.58)
Baseline controls	No	Yes	Yes	No	Yes	Yes
Further controls	No	No	Yes	No	No	Yes
N observations	1203	1203	1188	1203	1203	1188
(Pseudo)R^2	0.07	0.40	0.45	0.02	0.14	0.17

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.
Note: OLS regressions use robust standard errors. High mean is a binary variable with 0 = low and 1 = high mean. Variance is a categorical variable with 0 = low variance, 1 = high variance and 2 = u-shaped. All baseline controls can take values between 0 and 4 token. Further controls include age, gender, education, acceptance of risk, trust, sucker aversion, free-rider aversion and measures for negative and positive reciprocity. All regressions control for order effects.

Somewhat unexpectedly, personal values in part II are also significantly higher in high mean conditions ($p < 0.01$), suggesting that they may be affected by information about the behavior of others, although the effect is smaller than for contributions (0.4 token). When having a closer look at personal values, we see that still about 70% of participants do not change their personal values. Most of the people who do change shift their personal values downwards ($p < 0.01$) and stated "middle" personal values in part I (1 or 3 token).[14] Those might be less salient contribution levels and could constitute a compromise between contributing nothing and half or everything and half. This may indicate a greater uncertainty about what is appropriate in the PGG and therefore make these participants more susceptible to be influenced by external information.

[14]See appendix A, Figure A.2 and A.3. In high mean conditions, participants who shift the most are those who stated personal values of 1 in part I (upward shift). In low mean conditions it is participants who stated personal values of 3 (downward shift).

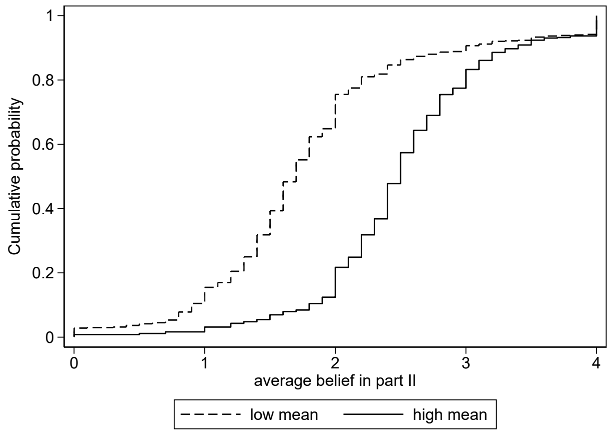


Figure 8: Average beliefs about other player in part II

5.3 Hypothesis 2: Effect of variance and shape of a distribution

We now address our main research question, namely how the variance and polarization of the norm affects the variance of individual responses. Recall that under a quadratic loss function, we only expect the mean of a distribution to matter (see *proposition 1*). By contrast, if we assume a linear-kinked loss function, we expect participants to also react to the variance such that loose norms lead to a larger variance in behavior. This also implies that a polarized norm in turn generates polarized behavior.

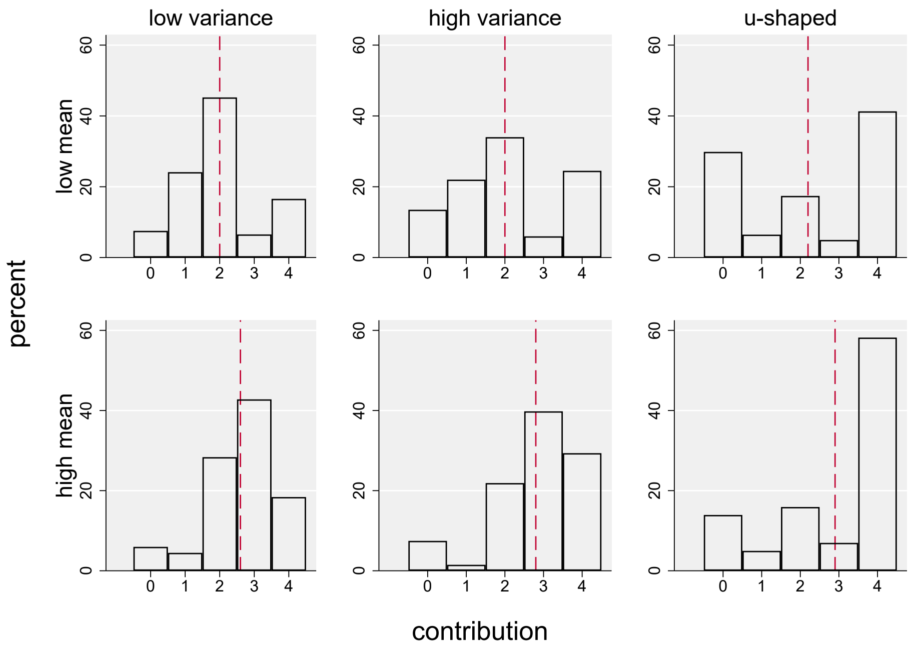


Figure 9: Distribution of contributions in part II by treatment

Figure 9 shows the distribution of contributions in part II for each experimental condition. We can see that there is a stark difference between distributions. In particular, we see that a tight norm (low variance condition) leads participants to choose contribution levels that are quite tightly centered around the mean of the shown distribution. Under loose (high variance) and polarized norms, by contrast, we see a much larger variation in behavior. In other words, loose norms generate loose responses while tight norms generate tight responses. What's more, we find that a polarized norm in turn causes polarized behavior.[15] Our data suggests that reciprocity concerns in the PGG can be better modelled by linear-kinked than quadratic loss functions.

Result 2: *Looser norms lead to a larger variance in behavior. When norms are polarized (U-shaped) they in turn generate polarized behavior.*

To test this first visual impression, we perform pairwise F-tests for the equality of standard deviations between distributions. Overall, we find that the variance in contributions is significantly higher in the u-shaped conditions than in the high or low variance conditions (both $F < 0.01$). The high variance conditions in turn have a significantly higher variance than the low variance conditions ($F < 0.01$).[16] As a robustness check we also test whether the distributions are statistically different using a Kolmogorov-Smirnov test.[17] The results confirm that overall contributions in the u-shaped condition are distributed significantly different than in the two other conditions (for both $p < 0.01$). Likewise, contributions in the low and high variance conditions follow a significantly different distribution ($p < 0.1$).[18]

When looking at the participants' beliefs about their co-player's contribution, we see that these largely mirror the distribution they were given in the sense that a larger variance in observed behavior leads to a larger variance in beliefs. This again suggests that a main channel through which the provided information affects contribution behavior is through shifting beliefs.[19] In addition, we also see that the variance of personal values is affected by the different treatment

[15] Reactions to a polarized norm are thus quite heterogeneous. See section 5.5 for a discussion about individual factors that could influence who focuses on which tail of the distribution.

[16] When looking separately at high and low mean conditions, the u-shaped distributions always show significantly higher variances ($F < 0.01$). The difference between high and low variance conditions is significant for the low mean conditions ($F < 0.05$), but not for the high mean conditions.

[17] See appendix A, Figure A.1 to see the cumulative distributions across treatment conditions.

[18] Again, when looking separately at high and low mean conditions, the result is robust for the u-shaped conditions. However, for high and low variance conditions we do not find a significant difference.

[19] See Appendix A Figure A.4/A.5 for how (normative) beliefs are distributed in the second part of the experiment.

conditions. Personal values vary more in the u-shaped condition than in either low or high variance conditions (F test, $p < 0.01$). Moreover, personal values in high variance conditions vary more than personal values in low variance conditions (F test, $p < 0.01$).

5.4 Hypothesis 3: importance of personal values across norm environments

Our final hypothesis states that under a linear-kinked loss function we expect personal values to have a larger effect in loose and polarized than in tight norm environments. Intuitively, if social norms are loose or polarized and people observe a larger behavioral variance, personal values could play a larger role in guiding decisions due to the higher strategic uncertainty. To test this hypothesis formally, we run the following regression:

$$C_i = \beta_0 + \beta_1 PV_i + \beta_2 Variance_i + \beta_3 PV_i \text{ x } Variance_i + \beta_4 Mean_i (+\beta_5 X_i) + \epsilon_i, \qquad (5)$$

where C_i is again the number of token an individual contributes in part II of the experiment, $Variance_i$ and $Mean_i$ are treatment indicators for our different distributions and PV_i captures the personal value an individual stated in part I before being confronted with any distributions. X_i is a vector of individual controls.

As can be seen from Table 4 the interaction between personal values and the high variance conditions is positive and significant across all specifications. Similarly, the interaction between personal values and the u-shaped conditions is always positive and significant in all but one condition.[20] This means our data confirms that the effect of personal values on contribution behavior is stronger in loose and polarized than in tight norm environments.

Result 3: *Personal values matter more for individual behavior when norms are loose or polarized compared to when they are tight.*

5.5 Individual determinants of different contribution levels

In addition to our treatment conditions, there is a number of factors that could affect contributions. Table 5 shows the results from regressing contributions in part II on demographics and individual characteristics. While demographics do not seem to play a role, we see a consistent effect of several individual characteristics. Firstly, holding everything else equal participants with

[20]In column (1) of Table 4 the p-value for the coefficient on u-shaped x personal values in part I is $p = 0.11$.

Table 4: Effect of personal values across different variances. Dependent variable = contribution in part II

	OLS		Tobit	
	(1)	(2)	(3)	(4)
personal values in part I	0.39***	0.31***	0.60***	0.48***
	(0.05)	(0.05)	(0.09)	(0.09)
Variance (baseline = low)				
high variance	-0.22	-0.21	-0.58	-0.60
	(0.20)	(0.19)	(0.39)	(0.37)
u-shaped	-0.13	-0.20	-0.55	-0.68*
	(0.25)	(0.24)	(0.41)	(0.39)
Interactions				
high variance × personal value in part I	0.13*	0.13*	0.30**	0.30**
	(0.07)	(0.07)	(0.13)	(0.12)
u-shaped × personal value in part I	0.13	0.15*	0.42***	0.45***
	(0.08)	(0.08)	(0.14)	(0.13)
high mean	0.73***	0.67***	1.06***	0.97***
	(0.07)	(0.07)	(0.12)	(0.12)
Constant	0.86***	0.17	0.30	-1.25*
	(0.13)	(0.44)	(0.29)	(0.68)
Further controls	No	Yes	No	Yes
N observations	1203	1188	1203	1188
(Pseudo) R^2	0.23	0.32	0.07	0.11

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.
Note: OLS regressions use robust standard errors. High mean is a binary variable with 0 = low and 1 = high mean. Variance is a categorical variable with 0 = low variance, 1 = high variance and 2 = u-shaped. Personal values in part I can take values between 0 and 4 token. Further controls include age, gender, education, acceptance of risk, trust, sucker aversion, free-rider aversion and measures for negative and positive reciprocity. All regressions control for order effects.

a higher acceptance of risk contribute significantly more to the public good. Moreover, when not controlling for baseline measures, higher trust in others correlates with higher contributions. Finally, our measures for *sucker* and *free-rider aversion* that proxy a participant's aversion of contributing too much/ too little are highly significant in all specifications, which underlines the strategic uncertainty inherent in the PGG.

Intuitively, the more people would be upset when finding out that they contributed everything and their co-player nothing (*sucker aversion*), the lower their contributions. By contrast, the more people would feel ashamed when contributing nothing and finding out that their co-player contributed everything (*free-rider aversion*), the higher their contributions.[21]

[21] We find that certain demographics correlate with free-rider and sucker aversion. Women seem to be more concerned with both types of aversions. Older people are less concerned with them. See appendix A Table A.3.

Table 5: Determinants of contribution decisions. Dependent variable = contribution in part II

	OLS		Tobit	
	(1)	(2)	(3)	(4)
Demographics				
age	0.00	0.00*	0.01*	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)
gender (female = 1)	-0.05	0.03	-0.10	0.05
	(0.08)	(0.07)	(0.14)	(0.11)
education (baseline = no formal degree)				
secondary school	0.10	-0.07	0.24	-0.10
	(0.29)	(0.27)	(0.45)	(0.37)
university/ college	-0.02	-0.20	0.02	-0.33
	(0.29)	(0.27)	(0.44)	(0.36)
prefer not to say	0.09	-0.13	0.02	-0.45
	(0.46)	(0.32)	(0.76)	(0.62)
Individual characteristics				
acceptance of risk	0.08***	0.05***	0.14***	0.09***
	(0.01)	(0.01)	(0.02)	(0.02)
trust	0.07*	-0.01	0.15**	0.01
	(0.04)	(0.03)	(0.07)	(0.06)
sucker aversion	-0.15***	-0.09***	-0.28***	-0.17***
	(0.02)	(0.02)	(0.04)	(0.03)
free-rider aversion	0.19***	0.11***	0.34***	0.19***
	(0.02)	(0.02)	(0.03)	(0.03)
positive reciprocity	0.06	0.02	0.13*	0.06
	(0.04)	(0.03)	(0.07)	(0.06)
negative reciprocity	-0.04	-0.03	-0.05	-0.03
	(0.03)	(0.02)	(0.04)	(0.04)
Constant	0.91**	-0.06	-0.17	-1.75***
	(0.44)	(0.39)	(0.68)	(0.58)
Observations	1188	1188	1188	1188
Baseline controls	No	Yes	No	Yes
(Pseudo)R^2	0.21	0.45	0.06	0.17

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.
Note: OLS regressions use robust standard errors. All columns include treatment indicators for mean and variance of the distribution. Baseline controls include contributions, average beliefs and personal values in part I. All regressions control for order effects.

When exploring individual determinants of contributions, a particularly interesting case is the u-shaped distribution, as responses to it are highly divergent. In section 5.3 we saw that personal values play a particularly important role in environments of larger strategic uncertainty. When looking at *sucker* and *free-rider aversion*, we find a similar pattern: While both considerations matter in all treatments, their effect is significantly larger in the u-shaped compared to the low variance condition (see appendix A, Table A.4).

6 Conclusion

In this study, we explore how the variance of norms affects individual behavior. We first develop a theoretical framework that is based on the assumption that players are motivated by reciprocity and interpret differences in the co-player's distribution as a shift in strategic uncertainty. We then test our framework empirically in the context of a PGG. To do so, we measure behavior in the PGG both before and after participants receive information about the distribution from which a co-players contribution is drawn. We thereby vary both the *mean* (high/low) and the *variance* (high/ low/ u-shaped) of the observed distribution.

Our results confirm previous research showing that information about average behavior has an important effect on subsequent decisions. Individuals contribute significantly more in high mean conditions than in low mean conditions. However, the mean is not the only important feature of the distribution. In line with our theoretical framework, we find that looser norms generate a larger variance in individual responses compared to tighter norms. In other words, "tight breeds tight" and "loose breeds loose". Moreover, we find that, when confronted with a polarized (U-shaped) distribution, participants' responses are themselves polarized. A possible interpretation of these results is that people have heterogeneous reactions to situations characterized by high strategic uncertainty, while they react rather similarly when strategic uncertainty is low. Finally, we find that personal values have a higher predictive power for contribution decisions in loose and polarized environments compared to looser norms. This suggests that an individual's reaction to strategic uncertainty may be mediated by their personal values. This in turn has practical implications for behavioral change interventions. For example, when intervening in contexts that have loose or polarized cultural norms, it may be more fruitful to focus on personal values, whereas when intervening in contexts that have tight cultural norms, it may be more fruitful to focus on the behaviors of others.

Overall, we show that when studying social norms it is crucial to not only consider the mean, but the whole distribution. Doing so provides a lot of analytical richness that can form the basis for a better understanding of the different behavioral patterns observed across societies.

References

Artinger, F., Exadaktylos, F., Koppel, H., and Sääksvuori, L. (2010). Applying quadratic scoring rule transparently in multiple choice settings: A note. Technical report, Jena Economic Research Papers.

Bašić, Z. and Verrina, E. (2020). Personal norms—and not only social norms—shape economic behavior. *MPI Collective Goods Discussion Paper,(2020/25).*

Bénabou, R. and Tirole, J. (2016). Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives*, 30(3):141–64.

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

Bicchieri, C. (2016). *Norms in the wild: How to diagnose, measure, and change social norms.* Oxford University Press.

Bicchieri, C. and Chavez, A. (2010). Behaving as expected: Public information and fairness norms. *Journal of Behavioral Decision Making*, 23(2):161–178.

Bicchieri, C. and Dimant, E. (2019). Nudging with care: The risks and benefits of social information. *Public Choice*, pages 1–22.

Bicchieri, C., Dimant, E., Gächter, S., and Nosenzo, D. (2022). Social proximity and the erosion of norm compliance. *Games and Economic Behavior*, 132:59–72.

Bicchieri, C., Dimant, E., and Sonderegger, S. (2020). It's not a lie if you believe the norm does not apply: Conditional norm-following with strategic beliefs. Working Paper Available at SSRN: https://dx.doi.org/10.2139/ssrn.3326146.

Bicchieri, C. and Xiao, E. (2009). Do the right thing: but only if others do so. *Journal of Behavioral Decision Making*, 22(2):191–208.

Bowles, S. and Gintis, H. (2013). *A Cooperative Species: Human reciprocity and its evolution.* Princeton University Press.

Bursztyn, L., Egorov, G., and Fiorin, S. (2020). From extreme to mainstream: The erosion of social norms. *American Economic Review*, 110(11):3522–48.

Chaudhuri, A., Paichayontvijit, T., et al. (2006). Conditional cooperation and voluntary contributions to a public good. *Economics Bulletin*, 3(8):1–14.

Cialdini, R. B. and Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In Gilbert, D. T., Fiske, S. T., and Lindzey, G., editors, *The Handbook of Social Psychology*, pages 151–192. New York: McGraw-Hill.

Ciranka, S. and van den Bos, W. (2020). A bayesian model of social influence under risk and uncertainty. *PsyArXiv*.

d'Adda, G., Dufwenberg, M., Passarelli, F., and Tabellini, G. (2020). Social norms with private values: Theory and experiments. *Games and Economic Behavior*, 124:288–304.

Dimant, E. (2019). Contagion of pro-and anti-social behavior among peers and the role of social proximity. *Journal of Economic Psychology*, 73:66–88.

Dimant, E. (2021). Hate trumps love: The impact of political polarization on social preferences. Working Paper Available at SSRN: https://dx.doi.org/10.2139/ssrn.3680871.

Dimant, E. (2022). Putting the 'social' in the norm: A novel method to elicit the shape of social norms. Mimeo.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the european economic association*, 9(3):522–550.

Elster, A. and Gelfand, M. J. (2021). When guiding principles do not guide: The moderating effects of cultural tightness on value-behavior links. *Journal of Personality*, 89(2):325–337.

Falk, A., Becker, A., Dohmen, T. J., Huffman, D., and Sunde, U. (2016). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *IZA Discussion Paper 9674*.

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

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, 114(3):817–868.

Feldhaus, C., Sobotta, T., and Werner, P. (2019). Norm uncertainty and voluntary payments in the field. *Management Science*, 65(4):1855–1866.

Fiorina, M. P. and Abrams, S. J. (2008). Political polarization in the american public. *Annu. Rev. Polit. Sci.*, 11:563–588.

Fischbacher, U. and Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American economic review*, 100(1):541–56.

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters*, 71(3):397–404.

Fosgaard, T. R., Gårn Hansen, L., and Wengström, E. (2020). Norm compliance in an uncertain world. *IFRO Working Paper*.

Frey, B. S. and Meier, S. (2004). Social comparisons and pro-social behavior: Testing" conditional cooperation" in a field experiment. *American Economic Review*, 94(5):1717–1722.

Gächter, S., Kölle, F., and Quercia, S. (2017). Reciprocity and the tragedies of maintaining and providing the commons. *Nature Human Behaviour*, 1(9):650–656.

Gelfand, M., Li, R., Stamkou, E., Pieper, D., Denison, E., Fernandez, J., Choi, V. K., Chatman, J., Jackson, J. C., and Dimant, E. (2021). Persuading republicans and democrats to comply with mask wearing: An intervention tournament.

Gelfand, M. J. (2021). Cultural evolutionary adaptations to threat. Current directions in psychological science. In press.

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., et al. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, 332(6033):1100–1104.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., and McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, 91(2):73–78.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., and Westwood, S. J. (2019). The origins and consequences of affective polarization in the united states. *Annual Review of Political Science*, 22:129–146.

Iyengar, S. and Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, 59(3):690–707.

Kandel, E. and Lazear, E. P. (1992). Peer pressure and partnerships. *Journal of political Economy*, 100(4):801–817.

Kerr, N. L., Rumble, A. C., Park, E. S., Ouwerkerk, J. W., Parks, C. D., Gallucci, M., and Van Lange, P. A. (2009). "how many bad apples does it take to spoil the whole barrel?": Social exclusion and toleration for bad apples. *Journal of Experimental Social Psychology*, 45(4):603–613.

Krupka, E. and Weber, R. A. (2009). The focusing and informational effects of norms on pro-social behavior. *Journal of Economic Psychology*, 30(3):307–320.

Krupka, E. L. and Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3):495–524.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In Roth, A. and Kagel, J., editors, *Handbookof Experimental Economics*. Princeton University Press.

Machina, M. and Pratt, J. (1997). Increasing risk: some direct constructions. *Journal of Risk and Uncertainty*, 14(2):103–127.

McConnell, C., Margalit, Y., Malhotra, N., and Levendusky, M. (2018). The economic consequences of partisanship in a polarized era. *American Journal of Political Science*, 62(1):5–18.

Qualtrics (2005). Qualtrics software, Version [January 2021] of Qualtrics. Copyright ©[2021]. Qualtrics, Provo, Utha, USA. Available at: https://www.qualtrics.com.

Quentin, A. (2016). distBuilder v 1.1.0 (1.1.0). Zenodo. https://doi.org/10.5281/zenodo.166736.

Robbett, A. and Matthews, P. H. (2021). Polarization and group cooperation. Working paper.

Roos, P., Gelfand, M., Nau, D., and Lun, J. (2015). Societal threat and cultural variation in the strength of social norms: An evolutionary basis. *Organizational Behavior and Human Decision Processes*, 129:14–23.

Shang, J. and Croson, R. (2009). A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. *The Economic Journal*, 119(540):1422–1439.

Thöni, C. and Volk, S. (2018). Conditional cooperation: Review and refinement. *Economics Letters*, 171:37–40.

A Additional figures and analysis

A.1 Hypothesis 1: Effect of high and low means

As can be seen from Figure A.1, contributions in the high mean conditions are shifted towards the right for all different variances.

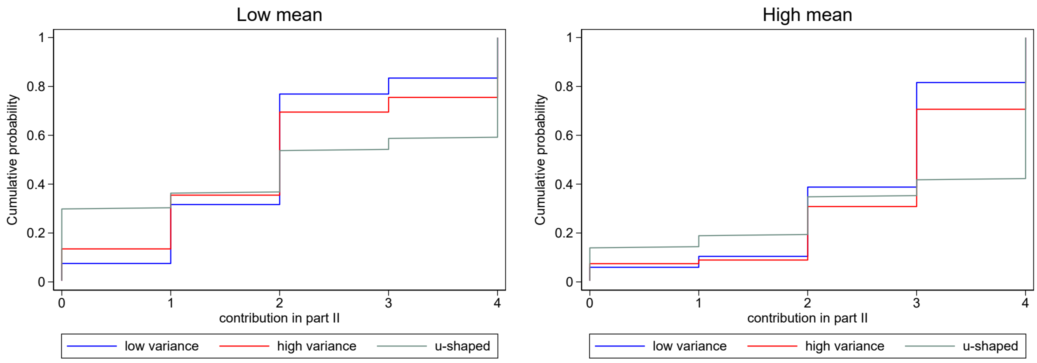


Figure A.1: Cumulative distribution of contributions in part II by treatment

Table A.1 is largely identical to the analysis we conduct in the main part of the paper. What it adds is an interaction between variance and mean conditions. As we can see, the OLS specifications show no significant interaction effect indicating that the effect of a high mean is identical across variances. When running a tobit regression, we see that the interaction becomes significant for the u-shaped distribution. There is thus some weak evidence that the effect of a high or low mean is stronger for polarized norms compared to a situation where norms are very tight. A potential reason for that could be that while u-shaped and low variance conditions have the same high and low means (2.4 vs 1.6 token), the modes of the distributions differ. In the u-shaped conditions a high mean implies a mode of 4 and a low mean a mode of 0. In the low variance conditions, by contrast, a high mean implies a mode of 3 and a low mean a mode of 1. In a way, the difference between high and low means is thus larger in the u-shaped conditions.

Table A.1: Effect of high and low mean conditions across different variances. Dependent variable = contribution in part II

	OLS			Tobit		
	(1)	(2)	(3)	(4)	(5)	(6)
high mean	0.63***	0.59***	0.54***	0.71***	0.64***	0.57***
	(0.11)	(0.09)	(0.09)	(0.23)	(0.18)	(0.17)
Variance (baseline = low)						
high variance	0.05	0.06	0.08	0.09	0.09	0.09
	(0.12)	(0.10)	(0.10)	(0.23)	(0.18)	(0.17)
u-shaped	0.21	0.20	0.19	0.29	0.28	0.24
	(0.14)	(0.13)	(0.12)	(0.23)	(0.18)	(0.18)
Interactions						
high mean x high variance	0.13	0.16	0.18	0.27	0.33	0.36
	(0.16)	(0.13)	(0.13)	(0.32)	(0.26)	(0.25)
high mean x u-shaped	0.06	0.17	0.17	0.64*	0.80***	0.78***
	(0.19)	(0.16)	(0.15)	(0.34)	(0.27)	(0.26)
Baseline controls						
contribution in part I		0.49***	0.45***		0.81***	0.75***
		(0.04)	(0.04)		(0.06)	(0.06)
belief in part I		0.19***	0.16***		0.34***	0.29***
		(0.05)	(0.05)		(0.08)	(0.08)
personal values in part I		0.04	0.03		0.06	0.05
		(0.04)	(0.04)		(0.06)	(0.06)
Constant	2.01***	0.21*	-0.00	2.12***	-0.93***	-1.55***
	(0.08)	(0.11)	(0.40)	(0.16)	(0.21)	(0.58)
Baseline controls	No	Yes	Yes	No	Yes	Yes
Further controls	No	No	Yes	No	No	Yes
N observations	1203	1203	1188	1203	1203	1188
(Pseudo)R^2	0.07	0.40	0.45	0.02	0.14	0.17

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.
Note: OLS regressions use robust standard errors. High mean is a binary variable with 0 = low and 1 = high mean. Variance is a categorical variable with 0 = low variance, 1 = high variance and 2 = u-shaped. Personal values in part I can take values between 0 and 4 token. Further controls include age, gender, education, acceptance of risk, trust, sucker aversion, free-rider aversion and measures for negative and positive reciprocity. All regressions control for order effects.

Figure A.2 and A.3 explore in more detail how personal values shift between part I and part II. As mentioned in our main analysis, we can see that a large share of participants (about 70%) does not change their personal values. If they do, we see that it is especially participants with "medium" initial personal values who change. In high mean conditions participants who stated a personal value of 1 are particularly likely to adjust upwards. Similarly, in the low mean conditions, it is participants who stated a personal value of 3 who are most likely to adjust downwards.

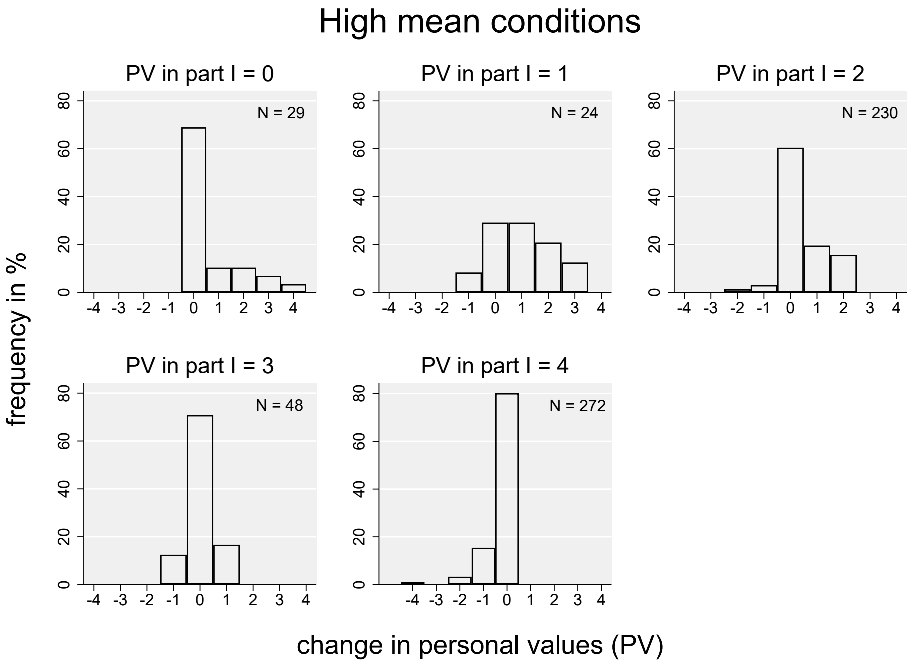


Figure A.2: Changes in personal values by initial personal values in part I

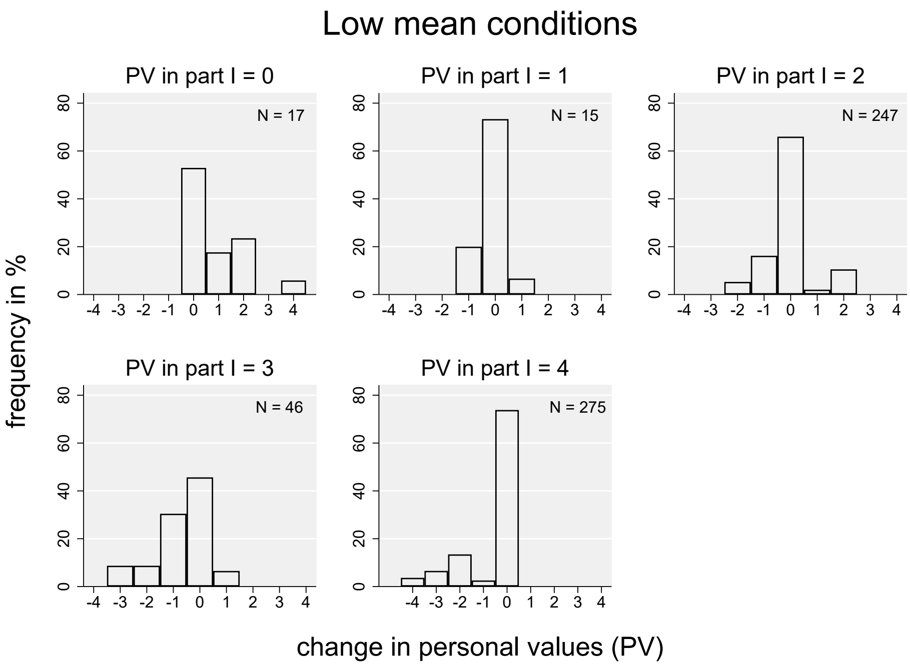


Figure A.3: Changes in personal values by initial personal values in part I

When observing that high and low mean conditions shift people's contribution behavior, a natural question is whether there are some people that are more likely to change their behavior. Table A.2 regresses changes in contributions (and personal values) on treatment indicators and individual characteristics. We find that compared to our tight norm condition (low variance), participants are less likely to change in loose and polarized environments. This is in line with the larger importance of personal values in these conditions. Apart from the treatment indicators, the only significant individual characteristics are sucker aversion and positive reciprocity. Intuitively, if individuals are highly averse to being taken advantage of or are very reciprocal they might react stronger to information about the co-player's behavior.

Table A.2: Likelihood of changing contribution and personal values in part II

	Changes in contributions	Changes in personal values
high mean	-0.01	-0.01
	(0.03)	(0.03)
Variance (baseline = low)		
high variance	-0.08**	-0.06*
	(0.04)	(0.03)
u-shaped	-0.07*	-0.02
	(0.03)	(0.03)
personal values in part I	-0.01	-0.06***
	(0.01)	(0.01)
Individual characteristics		
acceptance of risk	-0.00	0.00
	(0.01)	(0.01)
trust	-0.02	-0.02
	(0.02)	(0.02)
sucker aversion	0.03***	0.02***
	(0.01)	(0.01)
free-rider aversion	-0.00	-0.01*
	(0.01)	(0.01)
positive reciprocity	0.02	0.03**
	(0.02)	(0.01)
negative reciprocity	-0.01	-0.01
	(0.01)	(0.01)
Constant	0.33**	0.47***
	(0.15)	(0.15)
Observations	1188	1188
Demographic controls	Yes	Yes
R^2	0.04	0.05

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.
Note: OLS regressions using robust standard errors. Dependent variables are dummies that take the value 0 if participants did not change their contribution/ personal values between part I and II and 1 if they did. High mean is a binary variable with 0 = low and 1 = high mean. Variance is a categorical variable with 0 = low variance, 1 = high variance and 2 = u-shaped. Personal values in part I can take values between 0 and 4 token. Demographic controls are age, gender and education. None of them has an effect on the likelihood of change. All regressions control for order effects.

A.2 Hypothesis2: Effect of variance and shape of a distribution

Figures A.4 and A.5 show the distribution of beliefs about the co-player and of NEs in part II. As contributions, they are highly influenced by the observed distributions.

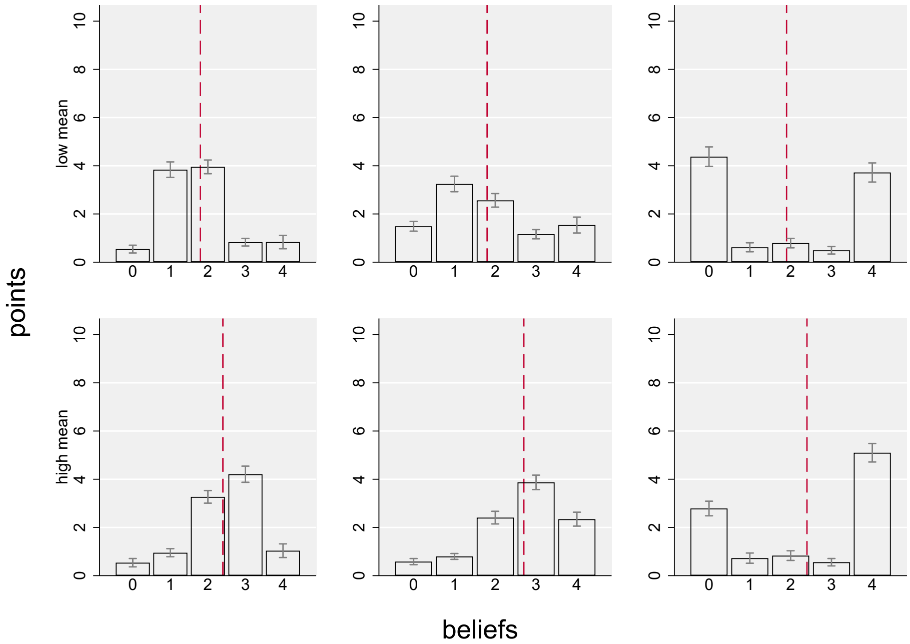


Figure A.4: Distribution of beliefs in part II by treatment

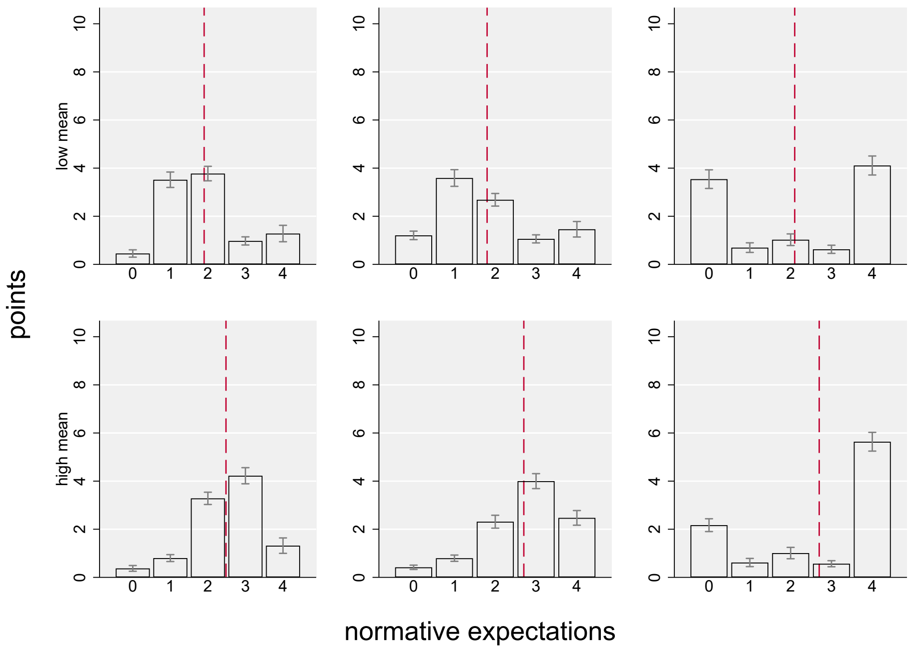


Figure A.5: Distribution of normative expectations in part II by treatment

A.3 Individual determinants of different contribution levels

In the main analysis, we found that sucker and free-rider aversion have a strong and significant effect on contribution behavior in the PGG. Table A.3 explores whether there are demographic variables that correlate with the two types of aversion. We find that women are significantly more likely to score high on both dimensions. By contrast, the older participants are the less they care about making very different contributions than the other player.

Table A.3: Determinants of free-rider and sucker aversion

	Sucker aversion		Free-rider aversion	
	(1)	(2)	(3)	(4)
gender (female = 1)	0.44***	0.74***	0.59***	1.15***
	(0.11)	(0.12)	(0.15)	(0.19)
age	-0.03***	-0.01***	-0.03***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.01)
education (baseline=no formal degree)				
secondary school	-0.31	-0.41	-0.63	-0.62
	(0.41)	(0.39)	(0.54)	(0.67)
university/ college	-0.39	-0.72*	-0.71	-1.13*
	(0.41)	(0.38)	(0.53)	(0.66)
prefer not to say	-0.13	-1.16*	-0.45	-1.67
	(0.50)	(0.60)	(0.93)	(1.18)
Constant	5.80***	4.82***	6.52***	5.13***
	(0.43)	(0.41)	(0.56)	(0.69)
N observations	1188	1188	1188	1188
(Pseudo) R^2	0.06	0.04	0.01	0.01

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.
Note: OLS regressions use robust standard errors. The dependent variables sucker and free-rider aversion are measured on a Likert scale from 1 to 7. (1) and (3) are OLS regressions, (2) and (4) are tobits, censored at 1 and 7. All regressions control for order effects.

Analogously to the differential effect of personal values across different variance conditions, we also find that sucker and free-rider aversion have a larger effect on contributions in polarized as compared to tight norm environments (see table A.4). When comparing high variance with low variance conditions, there is no significant difference although the signs of the coefficients go in the same direction as for the u-shaped conditions.

Table A.4: Effect of free-rider and sucker aversion on contributions across conditions. Dependent variable = contribution in part II

	OLS		Tobit	
	(1)	(2)	(3)	(4)
high mean	0.62***	0.67***	0.88***	0.95***
	(0.07)	(0.07)	(0.13)	(0.12)
Variance				
high variance	-0.11	0.13	-0.10	0.29
	(0.26)	(0.22)	(0.43)	(0.39)
u-shaped	0.14	0.20	0.50	0.62
	(0.27)	(0.25)	(0.43)	(0.40)
sucker aversion	-0.14***	-0.10***	-0.22***	-0.14**
	(0.03)	(0.03)	(0.06)	(0.06)
high variance × sucker aversion	0.01	-0.02	-0.01	-0.06
	(0.04)	(0.04)	(0.09)	(0.08)
u-shaped × sucker aversion	-0.09*	-0.10**	-0.25***	-0.28***
	(0.05)	(0.05)	(0.09)	(0.08)
free-rider aversion	0.15***	0.10***	0.23***	0.15***
	(0.02)	(0.02)	(0.05)	(0.05)
high variance × free-rider aversion	0.04	0.03	0.08	0.06
	(0.04)	(0.04)	(0.08)	(0.07)
u-shaped × free-rider aversion	0.11**	0.12***	0.28***	0.30***
	(0.04)	(0.04)	(0.08)	(0.07)
personal values in part I		0.42***		0.73***
		(0.03)		(0.05)
Constant	2.11***	0.84***	2.21***	0.04
	(0.17)	(0.18)	(0.30)	(0.32)
N observations	1203	1203	1203	1203
(Pseudo) R^2	0.18	0.30	0.05	0.09

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Note: OLS regressions use robust standard errors. High mean is a binary variable with 0 = low and 1 = high mean. Variance is a categorical variable with 0 = low variance, 1 = high variance and 2 = u-shaped. Personal values in part I can take values between 0 and 4 token. Sucker and free-rider aversion are measured on a Likert scale from 1 to 7. All regressions control for order effects.

B Instructions

Welcome

Thank you very much for participating in this study! This study consists of two parts and a questionnaire. Upon completion you will receive $1.70 for your participation plus an additional bonus of up to $2.36 that depends both on your decisions and the decisions of other participants. **In both parts you will face a situation in which you will be matched with one other, real participant**. On the next page we will describe this situation to you in more detail.

Instructions (1/2)

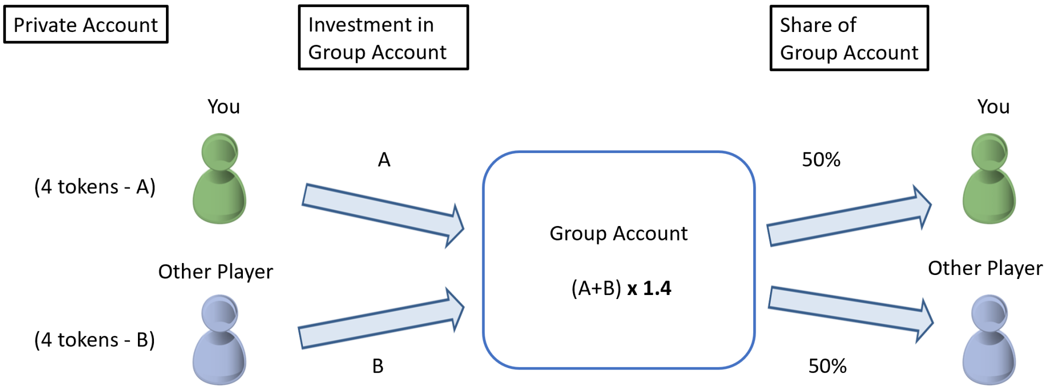
In this study, you will be anonymously paired with another participant. You will each start with **4 token** in your personal **private accounts**. In addition to the private accounts, there is a **group account**. You have to decide how many of your token you want to invest in the group account (either 0, 1, 2, 3 or 4 token). The amount leftover will remain in your private account. The other player has to make the same decision.

Your income from the private account

The amount in the private account is yours to keep. The other player doesn't earn anything from the token you keep in your private account. For example, if you keep 2 token in your private account, this will be your income from this account.

Your income from the group account

The amount invested in the group account will be multiplied by 1.4. That is, each token invested in the group account will yield 1.4 token for the group. The total amount in the group account will be split equally between you and your partner regardless of your individual investments. That is, each player receives half (50%) of the total amount in the group account.



If, for example, the sum of all investments in the group account by you and the other player is 6 token (A+B), then the group account yields 6 x 1.4 = 8.4 token. Both you and the other player would then receive 0.5 x 8.4 = 4.2 token from this account.

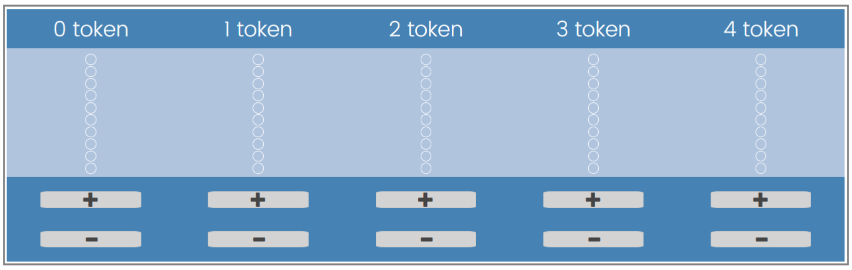
Your total income

Please note that for logistic reasons you are not interacting with other participants in real time. Once we collected all responses, we will match you with another person to calculate your and the other's total income. The latter consists of **all the token you kept in your private account plus half of the token** that you and that other participant invested **in the group account**. **Your total income will then determine your bonus payment**, with each token being worth $0.10. Whatever income you earned in token will be converted at this rate into actual money at the end of the experiment and paid out as a bonus.

Instructions (2/2)

In addition to making an investment decision in this situation, we will ask you to state **your beliefs about other participants**. You will be paid for these tasks according to how accurate your beliefs are.

A brief explanation follows: let us assume, we ask you to make a guess about how many token the participant you have been matched with invested in the group account. In this case, you would have to indicate **how likely you think it is that the other participant invested 0, 1, 2, 3 or 4 token**. To make your choices you will see a screen like the one below.



To make your decision you have to allocate a total of 10 points across options by clicking on the plus and minus buttons. The points you allocate need to add up to 10 and **the more likely you think one option is, the more points you would allocate to it**. The points you allocate to each option will naturally reflect your beliefs about the other participant's behavior.

The amount of money you can earn depends on how you allocated your points and what is actually true. If you put all points on the correct option, you will earn \$1 if you put all points on a wrong option you will earn \$0. In general, the more points you allocate to a correct option, the higher your earnings and the more points you allocate to a wrong option the lower your earnings. **The way your earnings are determined ensures that your best strategy is to carefully and honestly answer these questions.** If you want to have a closer look at how your earnings will be calculated click **here**.

Let's for example assume that you think it is equally likely that the other participant invested 2 or 4 token and you put 5 points on each option. If the other participant really invested either 2 or 4 token, you would in each case earn \$0.75. If they invested 0, 1 or 3 token you would earn \$0.

What if you had instead put all your eggs in one basket and allocated 10 points on the other participant investing 2 token? If the other participant indeed invested 2 token, you earn the maximum bonus of \$1. But if any of the other options is the correct one, you would earn nothing in this task. It is thus up to you to balance the strength of your personal beliefs with the risk of them being wrong.

In total, we will ask you to state your belief on **five** different questions throughout this study. In the end, a lottery will decide **one of them to be chosen for payment**. The amount you earned in the chosen question will then be added to your bonus payment.

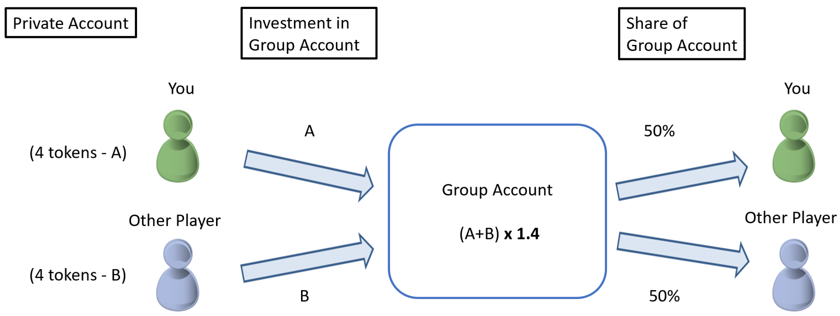
If participants click to get more information about payoffs, they see the following pop-up:

Your earnings are calculated on the basis of the table below. The more points you put on the correct option the higher your earnings. For each wrong option to which you allocate points your earnings will be reduced. The reduction is larger the more points you allocated to that option.

Points put on the correct option	**Earnings from the correct option**	**Points put on a wrong option**	**Costs from a wrong option**
10	100¢	10	50¢
9	99.5¢	9	40.5¢
8	98¢	8	32¢
7	95.5¢	7	24.5¢
6	92¢	6	18¢
5	87.5¢	5	12.5¢
4	82¢	4	8¢
3	75.5¢	3	4.5¢
2	68¢	2	2¢
1	59.5¢	1	0.5¢
0	50¢	0	0¢

Part 1

We are now going to ask you a number of questions that relate to the situation that you previously read (see image below). It is important that you answer these questions truthfully and as accurately as possible.[22]

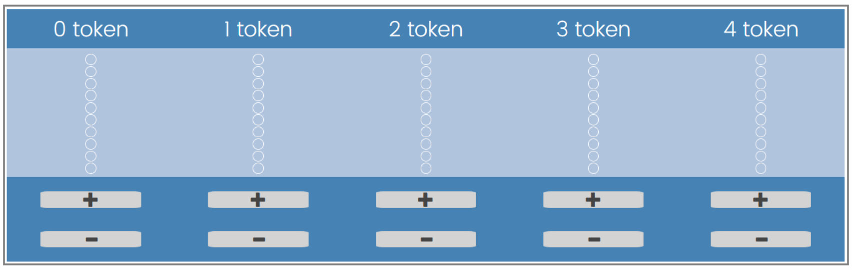


[22]Either normative questions or ABC questions are asked first. The three normative questions appear in randomized order.

1) We asked other participants what they believe is the **most appropriate amount to invest** in the **group account**. What do you believe was the most common answer?

Appropriate here means what you personally consider to be "correct" or "moral". You can allocate up to 10 points to an option by clicking on the plus and minus buttons below to indicate your guess. The more likely you think one option is, the more points you would allocate to it. Your guesses need to add up to 10.

Most people believe it is appropriate to invest...

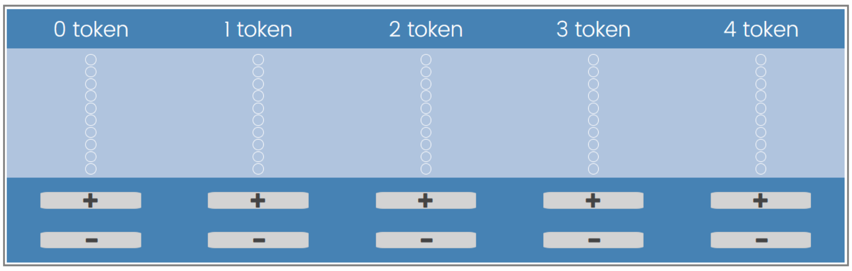


How confident are you in your response above? (0 not very confident, 100 very confident)

2) We asked other participants to make an investment decision in this situation. How many token do you believe most people **actually invested** in the group account?

Appropriate here means what you personally consider to be "correct" or "moral". You can allocate up to 10 points to an option by clicking on the plus and minus buttons below to indicate your guess. The more likely you think one option is, the more points you would allocate to it. Your guesses need to add up to 10.

Most people actually invested...



How confident are you in your response above? (0 not very confident, 100 very confident)

3) According to your own opinion and independent of the opinion of others, what is the **most appropriate amount to invest** in the **group account**?

Appropriate here means what you personally consider to be "correct" or "moral".

- 0 token
- 1 token
- 2 token
- 3 token
- 4 token

For the next section of Part 1, you will be matched with one other participant. You will only interact once with this person and you will never learn each other's identity.
Your and the other participant's bonus payment for Part 1 will depend on your decisions and the decisions of this participant.

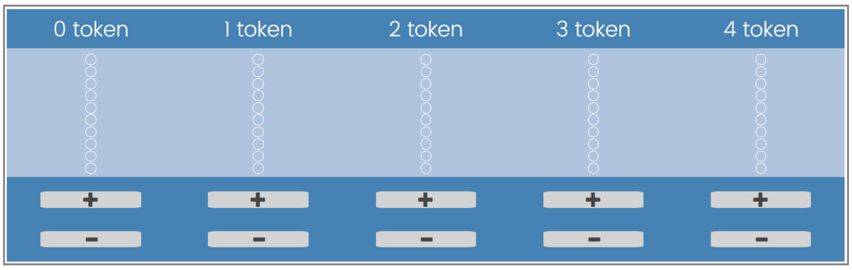
1) How many token do you want to **invest** in the **group account**?

- 0 token
- 1 token
- 2 token
- 3 token
- 4 token

2) How many token do you believe the participant you are matched with **invested** in the **group account**?

You can allocate up to 10 points to an option by clicking on the plus and minus buttons below to indicate your guess. The more likely you think one option is, the more points you would allocate to it. Your guesses need to add up to 10.

The participant you are matched with invested...



How confident are you in your response above? (0 not very confident, 100 very confident)

3) We are also interested in how many token you want to invest in the group account if you could know the other's choice beforehand. This means **you can condition** your investment on your group member's choice.

For one of you, the **unconditional choice** that you took before will count as the investment decision. For the other, the **conditional choice** (according to the table below) will count as the investment decision. Should the conditional choice be selected for you and the other participant invested x token in their unconditional choice, your decision for that scenario will determine your investment and thus matter for your bonus.

To determine your conditional choice, please tell us what you **want to invest** in the **group account** if:

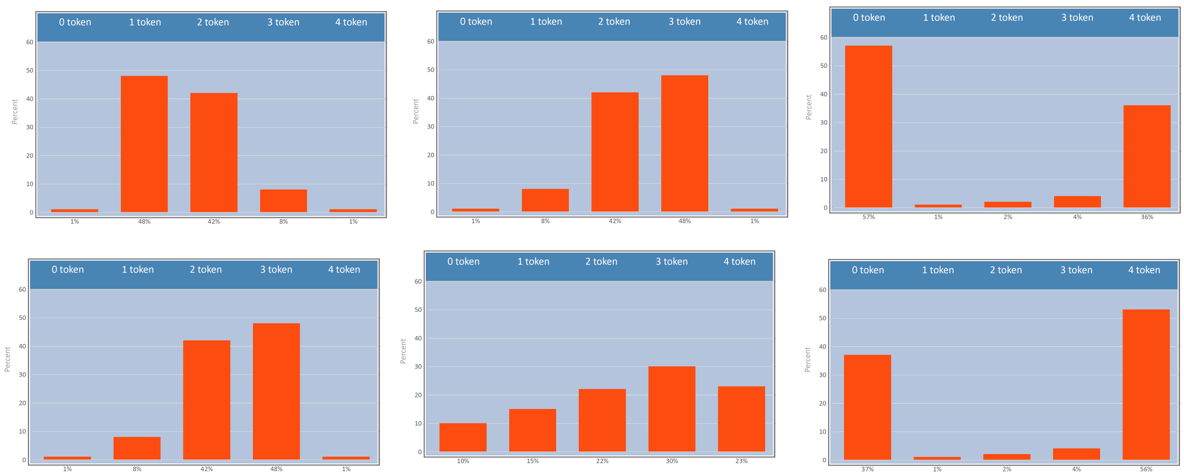
- The other player invests 0 token: ______ token
- The other player invests 1 token: ______ token
- The other player invests 2 token: ______ token
- The other player invests 3 token: ______ token
- The other player invests 4 token : ______ token

Part 2

In a previous study we asked over 600 participants to **make an investment decision** in the same situation. The possible choices were to invest 0, 1, 2, 3 or 4 token in the **group account**. From their answers we constructed different sub-groups. The graph below shows the percentage of people choosing each option in one randomly selected sub-group.

What previous participants invested in the group account:

(Participants are randomly shown one of the following six pictures.)



For Part 2 of the experiment you are matched with **one of the participants from the sub-group above**. You will only interact once with this person and you will never learn each other's identity.

Your and the other participant's **bonus payment for part 2** will depend on your decisions and the decisions of this participant.[23]

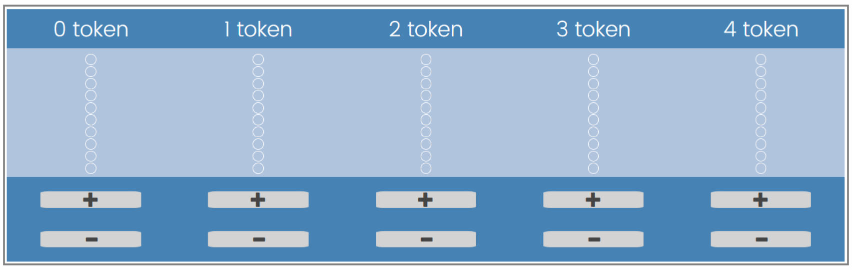
1) We asked other participants from the previous study what they believe is the **most appropriate amount to invest** in the **group account**. What do you believe was the most common answer?

Appropriate here means what you personally consider to be "correct" or "moral". You can allocate up to 10 points to an option by clicking on the plus and minus buttons below to indicate your guess. The more likely you

[23] We randomise whether participants are first asked about contributions and beliefs or about personal values and normative expectations.

think one option is, the more points you would allocate to it. Your guesses need to add up to 10.

Most people believe it is appropriate to invest...



How confident are you in your response above? (0 not very confident, 100 very confident)

2) Think again about the investment decision itself. According to your own opinion and independent of the opinion of others, what is the **most appropriate amount to invest** in the **group account**?

Appropriate here means what you personally consider to be "correct" or "moral".

- 0 token
- 1 token
- 2 token
- 3 token
- 4 token

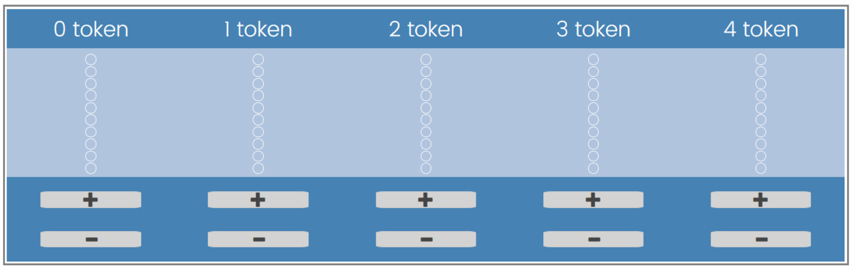
3) How many token do you want to **invest** in the **group account**?

- 0 token
- 1 token
- 2 token
- 3 token
- 4 token

4) How many token do you believe the participant you are matched with **invested** in the **group account**?

You can allocate up to 10 points to an option by clicking on the plus and minus buttons below to indicate your guess. The more likely you think one option is, the more points you would allocate to it. Your guesses need to add up to 10.

The participant you are matched with invested...

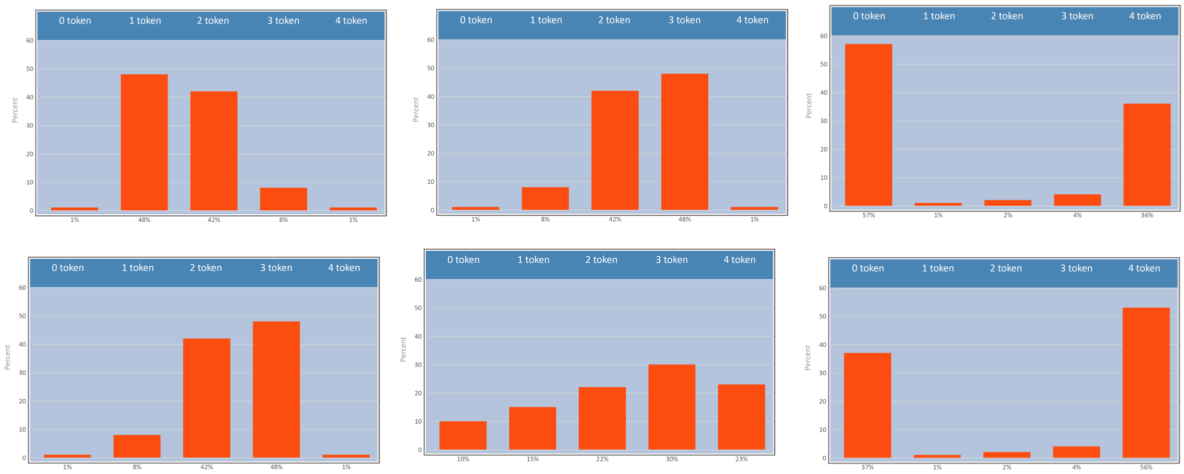


How confident are you in your response above? (0 not very confident, 100 very confident)

Questionnaire (1/2)

In this survey we showed you how many token a sub-group of other participants **invested** in the **group account**. Their answers are represented by the graph below.

What previous participants invested in the group account: *(Participants are randomly shown one of the following six pictures.)*



We will now ask you a few questions about the graph.

1) What are your thoughts on the behavior shown above?

____

2) Would you say the graph shows that overall

- people invest most of their token in the group account
- people invest half of their token in the group account
- people keep most of their token in their private account

3) Would you say the graph shows that overall

- there is a strong tendency for people to invest similar amounts in the group account
- there is a moderate tendency for people to invest similar amounts in the group account
- investments in the group account are very mixed

4) How common do you think the distribution of behavior shown above would be in other groups? (1 very rare, 7 very common)

5) How difficult was it for you to interpret the graph in Part 2, which is also shown above? (1 very easy, 7 very difficult)

6) How upset would you be if you invested everything in the group account and discovered that the participant you have been matched with invested nothing? (1 not at all upset, 7 very upset)

7) How ashamed would you be if you invested nothing in the group account and discovered that the participant you have been matched with invested everything? (1 not at all ashamed, 7 very ashamed)

Questionnaire (2/2)

1) Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?

- Need to be very careful
- Don't know
- Most people can be trusted

2) In how far do you agree with the following statement: "When someone does me a favour, I will return it." (1 don't agree at all, 7 completely agree)

3) In how far do you agree with the following statement: "If I am treated very unjustly, I will take revenge, even if there is a cost to do so." (1 don't agree at all, 7 completely agree)

4) Please tell me, in general, how willing or unwilling you are to take risks. (1 very unwilling to take risks, 11 very willing to take risks)

5) What is your age? ______ token

6) Which gender do you identify with?

- Female
- Male
- Non-binary
- Other
- Prefer not to say

7) What is the highest level of schooling you completed?

- No formal qualifications
- Secondary school
- University/ college degree
- Prefer not to say

Thanks a lot for participating in this survey! If you have any feedback for us you can write it here: