School of Economics Working Paper
2021-03

University of East Anglia

SCHOOL OF ECONOMICS

By Chance or by Choice? Biased Attribution of Others' Outcomes when Social Preferences Matter

Nisvan Erkal*
Lata Gangadharan**
Boon Han Koh***

*University of Melbourne, Australia
**Monash University, Australia
***University of East Anglia, Norwich

School of Economics
University of East Anglia
Norwich Research Park
Norwich NR4 7TJ
United Kingdom
www.uea.ac.uk/economics

By Chance or by Choice? Biased Attribution of Others' Outcomes when Social Preferences Matter*

Nisvan Erkal† Lata Gangadharan‡ Boon Han Koh§

March 2021

Abstract

Decision makers in positions of power often make unobserved choices under risk and uncertainty. In many cases, they face a trade-off between maximizing their own payoff and those of other individuals. What inferences are made in such instances about their choices when only outcomes are observable? We report findings from two experiments that investigate whether outcomes are attributed to luck or choices. We show that attribution biases exist in the evaluation of good outcomes. On average, good outcomes of decision makers are attributed more to luck as compared to bad outcomes. This asymmetry implies that decision makers get too little credit for their successes. Interestingly, the biases are exhibited by those individuals who make or would make the less prosocial choice for the group as decision makers, suggesting that a consensus effect may be shaping both the belief formation and updating processes.

JEL Classification: C92, D91, D81

Keywords: Decision-making under risk; Beliefs about others' decisions; Attribution biases; Social preferences; Consensus effect; Experiments

*We would like to thank the Editor and three anonymous reviewers, Kai Barron, Andy Brownback, Marina Gertsberg, Haley Harwell, Guy Mayraz, Vai-Lam Mui, Daniele Nosenzo, Silvia Sonderegger, Georg Weizsäcker, Tom Wilkening, seminar participants at WZB Berlin, University of Nottingham, University of Arkansas, University of Melbourne, Queensland University of Technology, Monash University, University of East Anglia, and Australian National University, as well as participants at the 11th Annual Australia New Zealand Workshop on Experimental Economics (ANZWEE), 86th Annual Conference of the Southern Economic Association (SEA), 8th Thurgau Experimental Economics Meeting, 2nd BEG Workshop on Behavioural and Experimental Economics (Gothenburg), 30th PhD Conference in Economics and Business, Behavioural Economics: Foundations and Applied Research Workshop (BEFAR), 2018 Asia Pacific Economic Science Association Conference, 2018 Economic Science Association World Meeting, and CPMD Workshop on Behavioural and Experimental Economics for their comments and feedback. We gratefully acknowledge funding from the Australian Research Council (DP1094676).

†Corresponding author. Department of Economics, University of Melbourne, VIC 3010, Australia. n.erkal@unimelb.edu.au.

‡Department of Economics, Monash University, VIC 3800, Australia. lata.gangadharan@monash.edu.

§School of Economics, University of East Anglia, NR4 7TJ, United Kingdom. b.koh@uea.ac.uk.

1 Introduction

In many environments, the determinants of outcomes are not observable. Decision makers make unobserved choices under risk and uncertainty, and outcomes are determined by a combination of their choices and luck. For instance, a firm's profits are determined by both the business strategies taken by its managers and the macroeconomic factors that are beyond their control. How are outcomes evaluated in such situations? Are there systematic biases in the attribution of outcomes to the decision makers' choices versus luck? Do they receive too little or too much credit?

We explore these questions in a leadership context where the choices decision makers make under risk affect their own payoffs as well as those of other individuals. Leadership is often about decision making for others and inherently involves assuming responsibility for the outcomes of others (Ertac and Gurdal, 2012; Edelson et al., 2018). In many cases, decision makers face a trade-off between maximizing their own payoff and those of other individuals. For example, society's growing demand for corporate social responsibility, defined as sacrificing firm profits for social interest, demonstrates how decision makers in positions of power are expected to engage in prosocial activity (Bénabou and Tirole, 2010).

We report findings from two experiments designed to examine how individuals who are affected by the choices of the decision makers form inferences about the decision makers. Our experimental design emphasizes the role social preferences play in leadership and aims to analyze the inferences formed about this important personality trait of decision makers. Uncovering biases in the attribution of outcomes is important for understanding the attitudes towards decision makers and the decision-making environment.[1]

In Experiment 1, individuals in their role as decision makers make an investment choice on behalf of their group. They choose between two investment options with binary outcomes. The outcome to the group depends on both the decision maker's choice, which

[1]For example, within the policy domain, redistribution decisions may be driven by beliefs about the determinants of income (Alesina and Angeletos, 2005; Rey-Biel et al., 2018) or self-serving attribution biases (Deffains et al., 2016). Misattribution of determinants have also been shown to affect consumer choice (Haggag et al., 2019).

is unobservable to the other group members, and luck. A high investment leads to a higher probability of the good outcome for the group but comes at a higher private cost to the decision maker. Hence, one can also think of the high investment decision as a costly effort choice made by the decision maker that increases the group's surplus at a personal cost. Consequently, decision makers' choices are affected by their social preferences. Using this design, we examine the group members' initial beliefs about the decision maker's type, and how these beliefs are updated after observing the outcome of the choice made by the decision maker.

In our analysis of belief updating, we examine three issues. First, we study biases in the way prior beliefs are treated in the updating process when group members form inferences about decision makers' prosocial preferences. More precisely, taking Bayes' rule as the benchmark, we ask whether group members suffer from base-rate neglect (i.e., put too little weight on their prior beliefs) or confirmatory bias (i.e., put too much weight on their prior beliefs) relative to a Bayesian.[2] Second, we examine whether, relative to a Bayesian, group members respond too little or too much to new information about the choice made by the decision maker. Responding too little to a good (bad) outcome, for example, would imply that they believe decision makers act selfishly (prosocially) and luck plays a bigger role in determining outcomes. Third, we explore whether group members treat good and bad outcomes asymmetrically. For example, observing a bigger response to bad outcomes (as compared to good outcomes) implies that members believe the decision maker's choice plays a larger role in determining bad outcomes than good outcomes. That is, when they see a bad outcome, they are likely to blame the decision maker for acting selfishly, but when they see a good outcome, they do not attribute it to the decision maker's prosociality.

In Experiment 1, we find that group members consistently suffer from base-rate neglect. This indicates, for example, that members who are initially more optimistic about the likelihood that the decision maker made a high investment decision tend to over-

[2]See, for example, Kahneman and Tversky (1973), Nisbett and Borgida (1975) and Tversky and Kahneman (1982) on base rate neglect, and Lord et al. (1979), Darley and Gross (1983), Plous (1991) and Rabin and Schrag (1999) on confirmatory bias.

update their beliefs about the decision maker's behavior when they observe a bad outcome. After accounting for base-rate neglect, we find that on average, members under-respond to good outcomes and attribute them more to luck as compared to a Bayesian. In contrast, their response to bad outcomes is similar to a Bayesian. This asymmetry implies that members on average attribute good outcomes more to luck and bad outcomes more to the decision maker's selfish choice. As a result, decision makers get too little credit for their successes.

Our results reveal that the biases that characterize the belief-updating process do not depend on the way the decision maker is selected. In line with the simple theoretical framework we consider, the appointment mechanism affects the initial beliefs formed about the decision maker's type, before the outcomes are observed. For example, members believe that a group-appointed decision maker is more likely to act in the group's interest as compared to a randomly appointed decision maker. However, once we control for the impact of the appointment mechanism on the initial beliefs, we find that the appointment mechanism has no additional impact on the updated beliefs.

A feature of our design in Experiment 1 is that the decision makers' choices and members' beliefs are elicited using the strategy method, where all individuals first make choices as decision makers before reporting their beliefs as group members. This allows us to examine the relationship between individuals' choices as decision makers and their attribution of the decision makers' outcomes as members. Using this design, we uncover that the asymmetry we identify in the evaluation of good and bad outcomes is driven by those individuals who make the less prosocial choice for the group. That is, those who make lower investment choices as decision makers are more likely to attribute others' good outcomes to luck. This suggests that a consensus effect may be at play as individuals use their own behavior as the basis for updating their beliefs about others (Ross et al., 1977; Marks and Miller, 1987; Dawes, 1989).[3]

[3]See Engelmann and Strobel (2000) for a discussion of the difference between the consensus effect and the false consensus effect. Ascribing of one's own motivation, feelings, and behavior onto other people is also referred to as "social projection" (see, e.g., Holmes, 1968; Ames, 2004). The psychological motivations underpinning social projection may be a desire to conform or feel connected to others, or a desire to justify one's own behavior. Social projection and the consensus effect may also be a heuristic that individuals use when making decisions in uncertain environments.

We explore this result further in Experiment 2, where participants no longer play both roles in the experiment. We are interested in investigating whether the biases we observe in Experiment 1 still exist when group members do not have experience making choices as decision makers.[4] Participants are informed at the beginning of the experiment whether they have been assigned as a decision maker or a group member. These roles do not change throughout the experiment.

To investigate whether different types (i.e., prosocial versus selfish individuals) form and update their beliefs differently, we ask group members to report their beliefs *before* asking them to indicate, hypothetically, what their investment decision would have been if they were the decision maker. This allows us to test whether there exists a correlation between each group member's type and their belief irrespective of whether they play both roles or just one. That is, by eliciting these hypothetical decisions after the belief-elicitation stage, we are still able to examine the relationship between individuals' effort choices as decision makers and their beliefs as members.

Interestingly, a correlation between the attribution of good outcomes to luck and what members would have chosen as decision makers also emerges in Experiment 2. That is, those members who are more likely to attribute good outcomes to luck are also more likely to state afterwards that they would have chosen low effort if they were placed in the position of the decision maker. This leads us to conclude that with and without the experience of acting as a decision maker, the same biases exist and seem to be driven by a consensus effect.

Our paper is related to three strands of the literature. First, our study substantially advances the research on attribution biases in beliefs in both economics and psychology. Studies in experimental economics have analyzed biases in beliefs and information processing by focusing mainly on ego-related beliefs, i.e., beliefs about one's own ability or physical attributes where one's ego can play a big role in shaping their beliefs (Eil and Rao, 2011; Ertac, 2011; Grossman and Owens, 2012; Möbius et al., 2014; Coutts,

[4]Experience in the decision-making process could allow members to put themselves into the shoes of the decision maker. Such perspective taking may influence their behavior and biases (see, e.g., Galinsky and Moskowitz, 2000; Todd et al., 2011; Lange et al., 2019).

2019). Both Eil and Rao (2011) and Möbius et al. (2014) find evidence of asymmetric updating, where agents are more responsive to good news than to bad news about their own performance in an IQ test or a beauty task. While Grossman and Owens (2012) find no evidence of asymmetry, Ertac (2011) and Coutts (2019) find that individuals tend to overweigh bad news.[5]

The related literature in psychology has mainly focused on self-serving biases in the attribution of own versus others' outcomes (see, e.g., Miller and Ross, 1975). Consistent with our findings, individuals tend to attribute others' good outcomes to exogenous factors (such as luck). In comparison, they are more likely to attribute their own good outcomes to endogenous factors (such as ability). Similarly, Pettigrew (1979) finds that good outcomes of out-group members are attributed to luck, but the opposite pattern emerges for in-group members.

Our novelty in relation to this strand of the literature in both economics and psychology is that we focus on the evaluation of others' outcomes in a context where decision making is shaped by social preferences. We show that good and bad outcomes are treated asymmetrically in this case also, and attribution biases exist in the case of good outcomes only. Moreover, our findings reveal that individuals' evaluation of others' prosociality tend to be correlated with their own behavior.

Second, our work is related to the literature on outcome biases, where researchers also find asymmetric evaluation of others' good and bad outcomes. However, it is assumed in this literature that all determinants of outcomes are fully observable. Despite this, good outcomes are treated more favorably than bad outcomes, suggesting that evaluators are biased by luck (see, e.g., Charness and Levine, 2007; Gurdal et al., 2013; Brownback and Kuhn, 2019). Our research extends this literature by considering the (arguably more common) setup where determinants of outcomes are not observable.

[5]Asymmetries in information processing have been analyzed and found in other domains as well. For example, Garrett and Sharot (2017) and Sunstein et al. (2017) show how good news and bad news are treated differently in the context of different life events and climate change, respectively. Evidence also suggests that individuals may treat new information that is in line with their convictions and new information that is against their expectations asymmetrically. For instance, Nyhan and Reifler (2010) analyze this in the context of political news and Sharot et al. (2011) study this with respect to health outcomes.

Finally, our paper is related to the literature which uses observational data to investigate how individuals respond to others' favorable and unfavorable outcomes in contexts such as redistribution (Alesina and Angeletos, 2005), CEO compensation (Bertrand and Mullainathan, 2001; Leone et al., 2006), political elections (Wolfers, 2007; Cole et al., 2012), medical referrals (Sarsons, 2019), and soccer (Gauriot and Page, 2019).[6] A puzzling finding in this literature is that individuals are punished or rewarded for outcomes that they cannot fully control. Attribution biases such as those we find in our experiments may be one explanation for this phenomenon. However, examining biases in belief updating using observational data is challenging given that reliable data on subjective beliefs is often unavailable. Our experimental setting gives us the opportunity to examine attribution biases in a controlled environment with an objective signal generating process.

2 Experiment 1

2.1 Experimental Design

Figure 1 presents an overview of the experiment. The main task in the experiment is the investment task, which we explain in Section 2.1.1. According to our theoretical framework, decisions in the investment task are shaped by the subjects' social preferences. Hence, to elicit their social preferences, subjects also play the dictator game in groups of two. Each subject is given 300 Experimental Currency Units (ECU) and asked to allocate this endowment between themselves and their matched partner. Both subjects within the pair make allocation decisions as the dictator. They are told that one of the decisions will be randomly chosen at the end of the experiment to determine the final allocation of the given endowment within each pair. Once subjects play the dictator game, they receive instructions for the investment task.[7]

[6] See also Palfrey and Wang (2012), who study the responsiveness of prices to signals in asset markets.
[7] The instructions can be found in Appendix A.1.

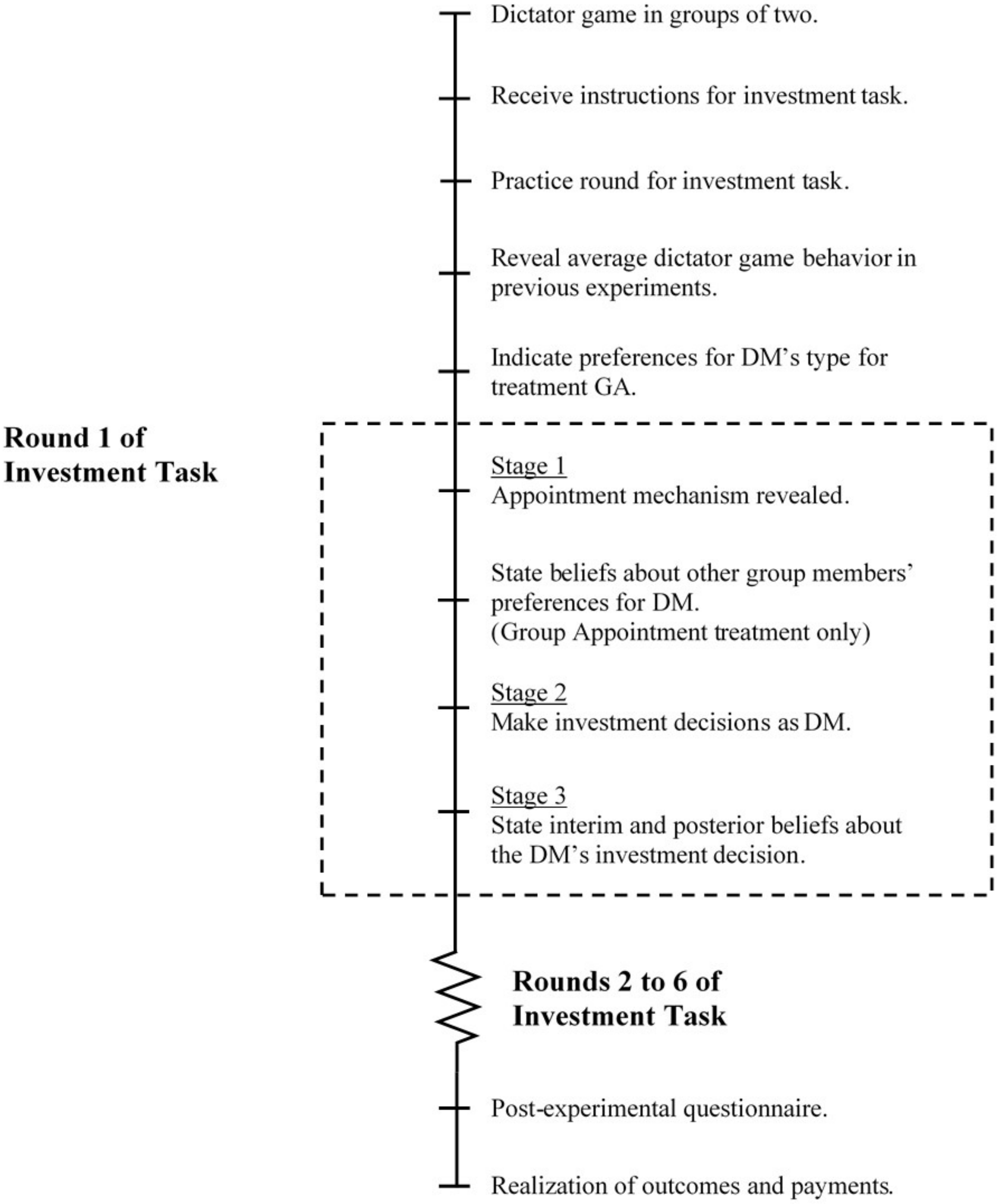


Figure 1: Overview of Experiment 1

2.1.1 Investment task

The experiment features a within-subject treatment design, where subjects play six repeated rounds of the investment task. In each round, subjects are re-matched to a new group with two other individuals (perfect stranger matching). Within each group, there is a decision maker (referred to as the DM in the rest of the paper) who makes an unobservable investment decision on behalf of the group. In the experiment, we label the DM as the leader.

Decisions are elicited using a strategy method. In each round, all subjects make their investment decisions assuming that they have been assigned to be the DM, and then state their beliefs about their DM's investment decision assuming that someone else in the group has been assigned to be the DM. This allows us to analyze whether beliefs are correlated with individuals' own decisions.[8] No feedback is given during the entire experiment. Subjects are informed whether they were assigned the role of the DM at the end of the experiment.

As shown in Figure 1, each round of the investment task consists of three stages, which we now explain in detail.

Stage 1: Appointment of DM. Group members' unconditional beliefs about their DM's investment decision can potentially depend on how the DM is appointed. We consider four mechanisms of appointing the DM. The appointment mechanism changes across the rounds, varying the initial beliefs members hold. This allows us to examine whether members' updating behavior depends on the distribution of their initial beliefs. For example, it may be the case that members are more likely to blame DMs for their failures if the DM is not appointed by the group.

At the beginning of each round, subjects are informed which mechanism will be employed in that round. In three of the appointment mechanisms, the DM is appointed exogenously. In the random assignment mechanism (treatment RA), each individual

[8]This setup resembles modern performance evaluation practices used by many organizations, such as 360-degree feedback, where decision makers are evaluated both by their colleagues (who may have similar responsibilities to them) and their subordinates (whose outcomes may be affected by the decision makers' actions) (see, e.g., Lepsinger and Lucia, 2009).

has an equal chance of being appointed as the DM. In the low and high assignment mechanisms (treatments LA and HA), subjects are informed that the group member who allocated the least and the highest amount to their matched partner in the dictator game, respectively, would be appointed as the DM.[9] Ties are broken randomly.

The fourth mechanism is the group appointment mechanism (treatment GA). Before beginning the first round of the investment task, each group member is asked to indicate whether they prefer: (i) to appoint the member who allocated the lowest amount to their matched partner in the dictator game; (ii) to appoint the member who allocated the highest amount to their matched partner in the dictator game; or (iii) to randomly select one member to be the DM. In addition, the subjects are asked to state their beliefs about the other two group members' preferences on which appointment mechanism to use.[10] To appoint the DM, the computer randomly picks one group member. This member's decision is used to determine which of the *other* two members will be the DM. This ensures that there is no scope for strategic behavior in that subjects are unable to influence their probability of being the DM through their decisions.[11] This is especially important in our set-up because, as explained later, there is a clear advantage to being the DM.

Stage 2: DM's investment decision. In the second stage of the investment task, each subject is asked to make an investment decision on behalf of the group.

The DM is given an individual endowment of 300 ECU and chooses between two investment options that will affect the payoffs of all the group members. The two investment options, given in Figure 2, are: (i) Investment X, which corresponds to a high effort level; and (ii) Investment Y, which corresponds to a low effort level. Both investment options yield the same high return if they succeed and the same low return if they fail. However, they differ in their probability of success/failure, and in their cost to the DM. Investment X succeeds with a probability of 0.75 and costs the DM 250 ECU, while In-

[9]Subjects are given the instructions for the investment task after they have made their decisions in the dictator game. This ensures that any strategic behavior in the dictator game is minimized. The actual decisions of subjects in the dictator game are not observed by their group members during the experiment. In the investment task, subjects are only informed about the mechanism that would be used to appoint the DM in each round.

[10]They are paid an additional 10 ECU if both of their guesses are correct.

[11]See, e.g., Galeotti and Zizzo (2018), for a similar protocol.

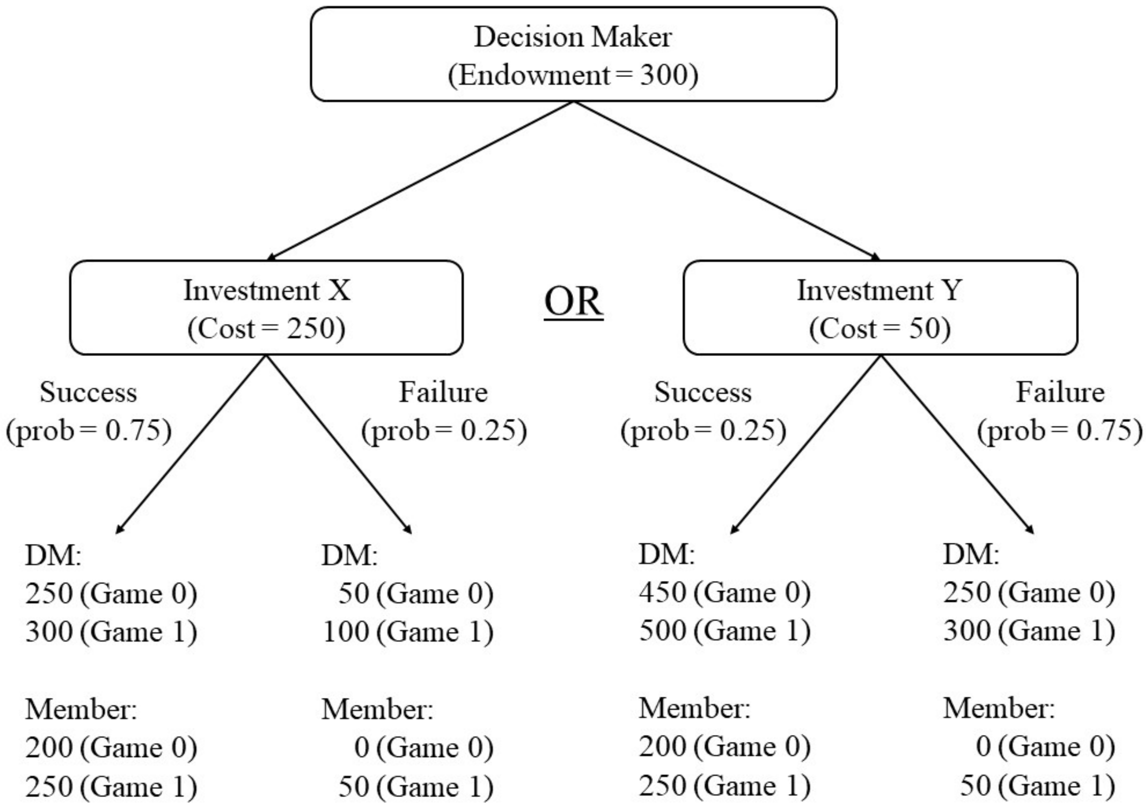


Figure 2: Investment task

vestment Y succeeds with a probability of 0.25 and costs the DM 50 ECU. Subjects are informed that the DM's investment decision will not be revealed to the group members. They only learn the outcome of the investment in the round randomly chosen for payment at the end of the experiment.

The returns from the two investment options are assumed to take the following values. In Game 1, the investment provides a return of 750 ECU for the group if it succeeds and 150 ECU if it fails. Note that the subjects' investment decisions as the DM and their beliefs as members about the DM's investment decision may be sensitive to the returns associated with the investment options. For instance, some subjects may be averse to choosing Investment Y if the members will receive a payoff of zero in case of failure. For this reason, we also consider Game 0, where the investment provides a return of 600 ECU if it succeeds and 0 ECU if it fails.

The return from the investment is distributed evenly between the DM and the two group members. The amount determines each group member's final payoff, except for the DM's. The DM's final payoff is equal to the sum of the endowment and the share of the return from the investment minus the cost of investment. These final payoffs are given at the bottom of Figure 2.

Stage 3: Elicitation of beliefs of group members. In the third stage of the investment task, subjects are asked to assume the role of group members and state their beliefs on the likelihood that the DM (i.e., one of the other two players in their group) has chosen Investment X. We elicit beliefs in the form of frequencies rather than probabilities.[12] When stating their beliefs, subjects are required to enter an integer number between 0 and 100.

We elicit two sets of beliefs from each subject.[13] First, each subject is asked to state their unconditional belief that the DM has chosen Investment X. Given that the subjects form these beliefs after being informed of the appointment mechanism, we refer to these unconditional beliefs as the members' *interim* beliefs. Second, each subject is asked to state their beliefs conditional on observing whether the investment chosen by the DM has succeeded or failed. We refer to these beliefs as the members' *posterior* beliefs. We do not impose any restrictions on their posterior beliefs. The group members can state any belief they want, regardless of what their interim beliefs are.

Subjects are paid for either their interim belief or their posterior beliefs. Beliefs are incentivized using the binarized scoring rule (BSR). We use the BSR because it incentives truth-telling independent of the subjects' risk preferences (Hossain and Okui, 2013; Erkal et al., 2020). It is a modified version of the quadratic scoring rule with a binary lottery procedure, where the distance between a subject's belief report and the DM's investment decision determines the probability of receiving a fixed amount (10 ECU in this case). As the subject's reported belief gets further away from the DM's investment decision, the

[12]Previous studies have found that subjects perform better in terms of Bayesian updating and additivity when beliefs are elicited as a population frequency. Gigerenzer and Hoffrage (1995), for example, find that subjects are more capable of performing Bayesian updating when probabilities are presented in the form of frequencies. Price (1998) finds that subjects are less likely to report extreme values in their beliefs when the questions are framed as relative frequencies. Schlag et al. (2015) argue that these findings point to the advantage of eliciting beliefs as frequencies rather than as probabilities.

[13]We used two separate screens. Screenshots of the decision screens can be found in Appendix B.

probability of receiving the fixed payment gets lower.[14]

2.1.2 Procedures and payment

The experiments were conducted in the Experimental Economics Laboratory at the University of Melbourne (E^2MU) and programmed using z-Tree (Fischbacher, 2007). We ran 10 sessions with 24 to 30 subjects in each session. A total of 282 participants, mostly students from the University of Melbourne, were recruited using ORSEE (Greiner, 2015).[15] Each session lasted between 90 and 120 minutes.

To ensure that the subjects fully understood the tasks, the experimenter verbally summarized the instructions after the subjects finished reading the printed instructions. Subjects completed a set of control questions and participated in a practice round using treatment GA and Game 0 before beginning the actual investment task. For Game 0, we implemented treatments LA and HA only since, as explained in the next section, theory suggests that the difference in interim beliefs should be the greatest between these two treatments. We implemented all four appointment mechanisms for Game 1, which allows us to study the subjects' behavior across different mechanisms using the same set of parameters.

The order between treatments was changed to control for potential order effects. However, since our main focus is the treatments associated with Game 1, Game 0 was always implemented in Rounds 1 and 2 while Game 1 was always implemented in Rounds 3 to 6. Table 1 summarizes the order of the treatments in each session. In each cell of the table, the first two letters denote the appointment mechanism, while the Arabic numeral at the end denotes the game faced by the subjects in the corresponding round within the session.[16]

[14]Specifically, for a given belief report $r \in [0, 100]$, the group member receives 10 ECU with probability $1 - \left[I(e = e_H) - \frac{r}{100}\right]^2$, where $I(e = e_H)$ is an indicator variable that equals 1 if the DM chose e_H (Investment X) and 0 otherwise.

[15]To avoid cultural factors influencing behaviour (Alesina and Angeletos, 2005; Rey-Biel et al., 2018), only Australian citizens were recruited. For eight subjects, we observed a mismatch between their citizen information on the recruitment system and their response relating to length of stay in the country in the post-experiment questionnaire. Further, two subjects had prior experience with the experiment. Hence, data from 272 subjects are used for the main analysis. Our results however are robust to the inclusion of the eight (non-Australian) subjects.

[16]For example, a cell that states "RA1" means that the DM was appointed randomly for that round

Table 1: Order of treatments for each experiment session

Session	# subjects	Round #						
		Practice	1	2	3	4	5	6
1, 5	60	GA0	LA0	HA0	RA1	LA1	HA1	GA1
2, 6	60	GA0	HA0	LA0	RA1	HA1	LA1	GA1
3, 8	54	GA0	LA0	HA0	GA1	LA1	HA1	RA1
4, 7	54	GA0	HA0	LA0	GA1	HA1	LA1	RA1
9, 10	54	GA0	LA0	HA0	LA1	HA1	GA1	RA1

At the end of the experiment, subjects were invited to complete a brief questionnaire which included demographic questions, questions about their decisions during the experiment, and an incentivized one-shot risk task (Gneezy and Potters, 1997) to elicit their risk preferences. Subjects were paid for either the dictator game or the investment task. If they were paid for the investment task, then we paid them for their decisions in one of the six rounds. For the chosen round, a DM was appointed according to the corresponding treatment and the DM was paid only for their investment decision. The other two members were paid for their DM's decision as well as their stated beliefs. Earnings were converted to cash at the conclusion of the session at the rate 10 ECU = 1 AUD. Overall, subjects earned between \$10 and \$76, with the mean earnings being \$34.07. Subjects' earnings also included a show-up fee of \$10.

2.2 Theoretical Framework

In this section, we provide a simple theoretical framework to evaluate how beliefs will be formed under the different appointment mechanisms.

2.2.1 DM's effort choice

Players maximize expected utility and are differentiated based on their other-regarding preferences. Let $\beta_i \in [0, 1]$ denote the type of player i. It is a private draw from a

(treatment RA), and the subjects played Game 1.

distribution $F(\beta)$ with density $f(\beta)$. $F(\beta)$ is common knowledge.[17]

Players are randomly assigned to groups of size $N > 2$. The DM in each group makes an effort choice $e \in \{e_L, e_H\}$ at cost $c \in \{c_L, c_H\}$ which is deducted from an initial endowment ω that the DM receives. Assume that $\omega \geq c_H > c_L > 0$. There are two possible team outputs, $Q \in \{Q_L, Q_H\}$, where $Q_H > Q_L$, and the DM's effort choice determines the probability with which each output level will be realized. A high effort choice leads to the high output level with a higher probability, but it costs more to the DM. Specifically, a high effort choice e_H leads to an output Q_H with probability p, where $p \in (0.5, 1)$, while a low effort choice e_L leads to an output Q_H with probability $1 - p$.

For a given outcome Q, each member in the group receives $\frac{Q}{N}$ and the utility of the DM is given by

$$U = u\left(\frac{Q}{N} + \omega - c\right) + \beta \cdot \sum_j v_j\left(\frac{Q}{N}\right). \tag{1}$$

$u(\cdot)$ and $v_j(\cdot)$ are twice differentiable utility functions with $u'(\cdot) > 0$ and $v_j'(\cdot) > 0$. $u(\cdot)$ represents the direct utility the DM receives from own monetary payoff while $v_j(\cdot)$ is the utility member j receives from own monetary payoff. β determines the weight the DM puts on the utilities of the other group members.[18]

DMs maximize their expected utility and choose e_H over e_L if $EU(e_H) \geq EU(e_L)$. In the experimental design, we refer to e_H and e_L as Investment X and Investment Y, respectively. The choice of parameters in Game 0 and Game 1 are $N = 3$, $\omega = 300$, $p = 0.75$, $Q_H = 750$ (Game 1) or 600 (Game 0), $Q_L = 150$ (Game 1) or 0 (Game 0), $c_H = 250$, and $c_L = 50$. Given these parameter choices, if $\beta = 0$, then the DMs only care about their own payoff and choose e_L since $EU(e_H) - EU(e_L) = p\big[u\big(\frac{Q_H}{N} + \omega - c_H\big) - u\big(\frac{Q_L}{N} + \omega - c_L\big)\big] + (1-p)\big[u\big(\frac{Q_L}{N} + \omega - c_H\big) - u\big(\frac{Q_H}{N} + \omega - c_L\big)\big] < 0$.[19] For $\beta > 0$,

[17]As indicated in Figure 1, subjects were informed that in previous experiments, (i) about 80% of participants transferred a positive amount to their matched partner, and (ii) for those who transferred, the average transfer was about 40% of their endowment. These statistics were obtained using data from pilot experiments ($N = 192$).

[18]See Rotemberg (2014) for a survey of models of social preferences used in the literature. Brock et al. (2013), Cappelen et al. (2013), and Exley (2015) analyze social preferences under risk. We differ from these papers with our focus on how the group members evaluate the outcomes of the DMs in risky environments.

[19]Note that $\frac{Q_H}{N} + \omega - c_L > \frac{Q_L}{N} + \omega - c_L = \frac{Q_H}{N} + \omega - c_H > \frac{Q_L}{N} + \omega - c_H$.

$EU(e_H) \geq EU(e_L)$ holds if

$$\beta \geq \beta^* \equiv \frac{\begin{bmatrix} p\left[u\left(\frac{Q_H}{N}+\omega-c_H\right)-u\left(\frac{Q_L}{N}+\omega-c_L\right)\right] \\ +(1-p)\left[u\left(\frac{Q_L}{N}+\omega-c_H\right)-u\left(\frac{Q_H}{N}+\omega-c_L\right)\right] \end{bmatrix}}{(1-2p)\sum_j\left[v_j\left(\frac{Q_H}{N}\right)-v_j\left(\frac{Q_L}{N}\right)\right]}.$$

Intuitively, DMs choose high effort if they care sufficiently about the payoffs of the other group members.[20] In the experiment, subjects' decisions in the dictator game provide a proxy for their types (β_i). We use the dictator game since it is widely used in the literature to measure social preferences.

2.2.2 Information and beliefs

Members' interim beliefs. We first consider the members' interim beliefs about their DM's type after observing the appointment mechanism. Specifically, we are interested in each member's belief that the DM is of type $\beta \geq \beta^*$, which corresponds to the likelihood that the DM chooses e_H over e_L. We denote member i's interim belief after observing appointment mechanism $\Psi \in \{RA, LA, HA, GA\}$ as μ_i^{Ψ}.

Our first testable prediction is about the ranking of the members' interim beliefs under the different appointment mechanisms:

Hypothesis 1: $\mu_i^{LA} \leq \mu_i^{RA} \leq \mu_i^{GA} \leq \mu_i^{HA}$.

The proof is in Appendix C. In treatment GA, all players prefer to have the highest type appointed as the DM. This is because all group members want the DM to choose e_H which maximizes their expected payoffs. Although this implies that the beliefs under treatments GA and HA should be the same, the difference stated in the hypothesis is due to the implementation strategy we follow in treatment GA. Specifically, the highest type in the group will not necessarily be appointed as the DM under treatment GA if his/her appointment decision is randomly picked to be implemented. Hence, $\mu_i^{GA} \leq \mu_i^{HA}$.

Members' posterior beliefs. We next consider how members update their beliefs about their DM's type after observing the outcome. The outcome $Q \in \{Q_L, Q_H\}$ is

[20] For instance, under the assumption of risk neutrality, $\beta^* = \frac{1}{2}$.

a signal that members receive about the DM's type. Note that $Pr(Q_L|\beta < \beta^*) = Pr(Q_H|\beta \geq \beta^*) = p$ and $Pr(Q_L|\beta \geq \beta^*) = Pr(Q_H|\beta < \beta^*) = 1 - p$.

We denote group member i's unbiased posterior belief of the DM's type, given a signal Q, as $\phi_i^{\Psi}|_Q$. Specifically, suppose the members receive a signal $Q = Q_H$. Using Bayes' rule, member i's posterior belief is given by

$$\phi_i^{\Psi}|_{Q_H} = \frac{\mu_i^{\Psi} \cdot Pr(Q_H|\beta \geq \beta^*)}{Pr(Q_H)} = \frac{\mu_i^{\Psi} p}{\mu_i^{\Psi} p + (1 - \mu_i^{\Psi})(1 - p)}.$$

$\phi_i^{\Psi}|_{Q_L}$ is defined in a similar way.

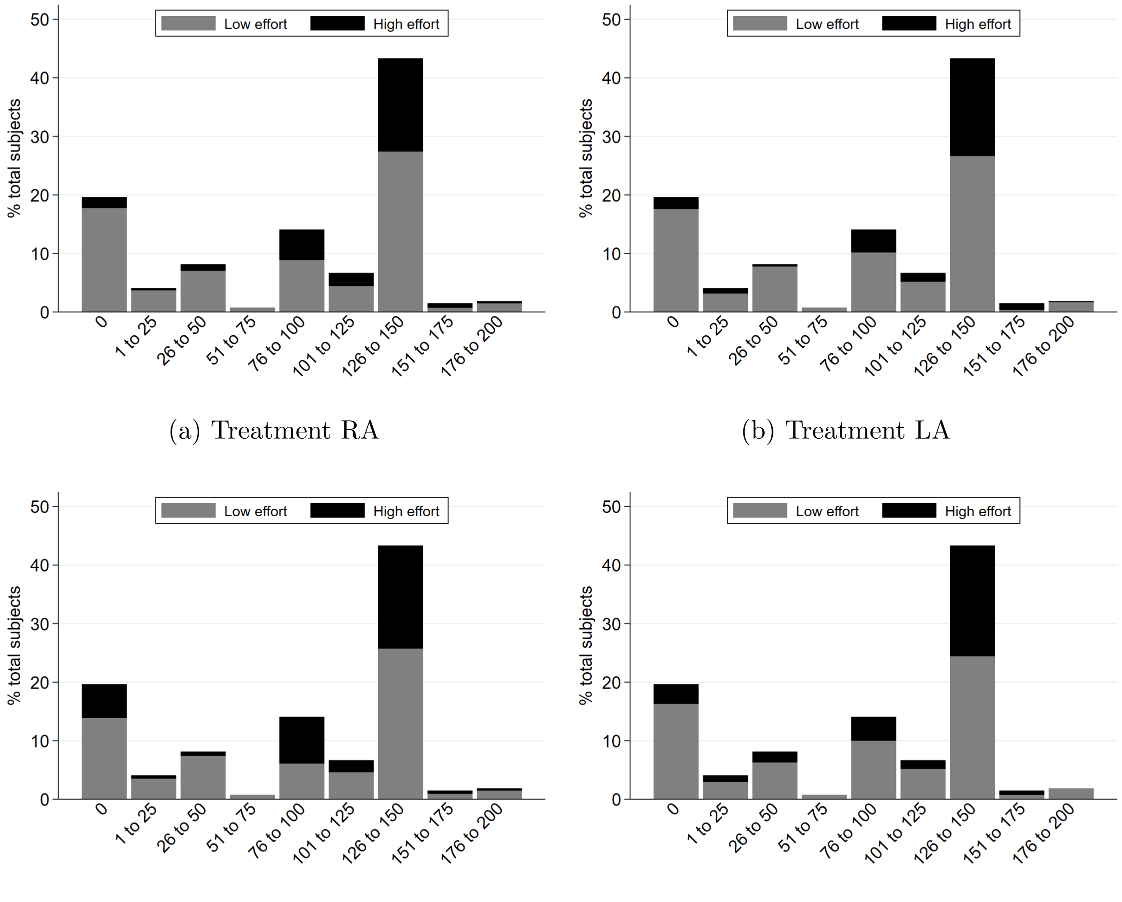
We test the null hypothesis that the members will be unbiased (i.e., Bayesian) when they update their beliefs:

Hypothesis 2: Group members behave like Bayesian agents when updating their beliefs about the DM under all of the appointment mechanisms.

We explain the econometric framework that we use to test Hypothesis 2 empirically in Section 2.3.3. If we detect deviations from the Bayesian benchmark (e.g., Tversky and Kahneman, 1974), our design allows us to investigate these deviations further in two ways. First, we can test whether the deviations vary across the different appointment mechanisms. That is, we can observe, for example, whether being appointed by the group (treatment GA) has an impact on the way members update their beliefs about the DM. Second, since we employ a strategy method, we can examine whether there is a correlation between subjects' decisions as DMs and their beliefs (initial and updated) about the DM. Such a correlation can be explained by a consensus effect. For example, a subject who chooses e_H may be more likely to believe that their DM has also chosen e_H.

2.3 Results

Since there were no interactions between the group members during the experiment and no feedback was given to the subjects from the previous rounds, our unit of observation is at the subject level. For the main analyses in this paper, we pool data from the Game



(a) Treatment RA

(b) Treatment LA

(c) Treatment HA

(d) Treatment GA

Figure 3: DM's effort choice in Game 1 against dictator game behavior

0 and Game 1 treatments.[21] For robustness, we show in Appendix D.1 that the main conclusions do not change when we consider the Game 1 treatments only.

2.3.1 The dictator game as a proxy for an individual's type

We conjecture in Section 2.2 that subjects' behavior in the dictator game is a proxy for their type. That is, subjects who transfer more of their endowment to their matched partner in the dictator game are more likely to choose e_H when they are in the role of the DM. As our testable hypotheses depend on this relationship between the DM's type and their effort choice, we first examine if it holds.

Figure 3 presents the distribution of subjects' decisions in the dictator game against their effort choices across different appointment mechanisms. Because the subjects only

[21]Hence, data from treatments LA0 and LA1 are pooled together as treatment LA, while data from treatments HA0 and HA1 are pooled together as treatment HA.

participate in the dictator game once, the distribution of transfers are the same across the different treatments. Within each panel in Figure 3, the black bars represent the proportion of DMs who choose high effort (e_H) while the gray bars represent the DMs who choose low effort (e_L).

A clear pattern that emerges is that DMs who are more prosocial in the dictator game are also the ones who are more likely to choose the investment option that is in the interest of the group (i.e., high effort). This pattern is consistent across the different appointment mechanisms.[22] The correlation between the DM's behavior in the dictator game and their effort choice are statistically significantly positive in all treatments. The Spearman's rank correlation coefficients and corresponding p-values are: (i) RA: 0.233, p-value < 0.001; (ii) LA: 0.262, p-value < 0.001; (iii) HA: 0.096, p-value $= 0.025$; and (iv) GA: 0.183, p-value $= 0.003$.

Table 2 presents marginal-effects estimates from a probit model for the relationship between the subjects' decisions as DMs in the investment task and their dictator game behavior. In the regression analysis, we control for order effects, the subjects' behavior in the risk task, the appointment mechanisms, and Game 1. We find a statistically significant and positive relationship between the DM's decision in the dictator game and their decision to choose high effort in the investment task (p-value < 0.001). A DM who transfers 1% more of their endowment to their matched partner in the dictator game is 0.4% more likely to choose e_H in the investment task on average. In addition, consistent with our expectations about the DM's behavior between the Game 0 and Game 1 treatments, we observe that subjects are 6.7% less likely to choose e_H in Game 1 on average, and this effect is statistically significant (p-value $= 0.002$).[23]

The established link between dictator game behavior and effort choices implies that

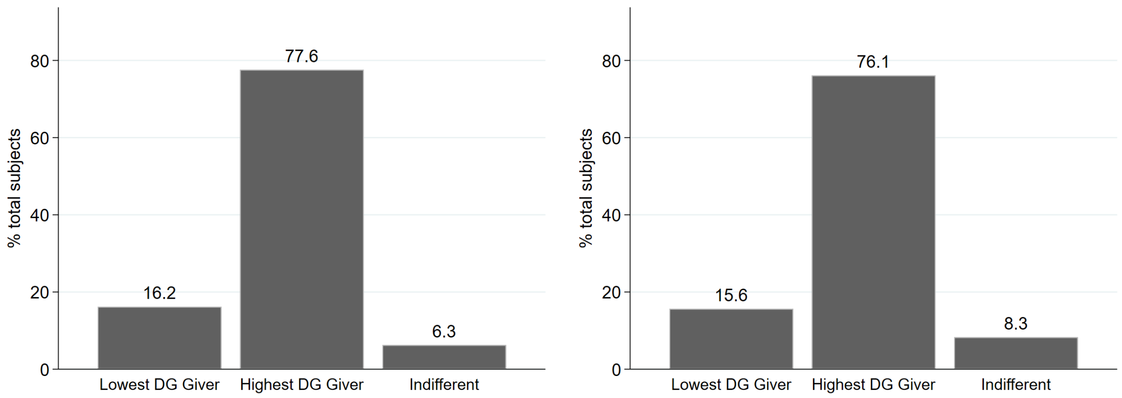
[22]Interestingly, there is a higher proportion of individuals who transfer nothing to their matched partner in the dictator game, but who choose high effort as DMs in treatment HA compared to the other treatments. This may be because these individuals believe that they are unlikely to be appointed as the DM in this treatment and therefore think that their effort choice is less likely to be implemented.

[23]We also elicited members' incentivized beliefs about their DM's behavior in the dictator game under each appointment mechanism. We find that there is a positive relationship between group members' interim beliefs and their reports of how much the DM has transferred in the dictator game. This further shows that the subjects regard the dictator game as a predictor of an individual's likelihood of choosing high effort as a DM.

Table 2: Regression of DM's effort choice

Variables	Dependent variable: =1 if DM chooses e_H (1)
% endowment transferred in DG	0.004***
	(0.001)
% endowment invested in RT	−0.001
	(0.001)
Treatment LA	−0.044*
	(0.026)
Treatment HA	0.046
	(0.029)
Treatment GA	0.040
	(0.029)
Game 1	−0.067***
	(0.022)
Order Effects	Y
Observations	1,632
# subjects (clusters)	272

Marginal effects of probit model reported. Robust standard errors in parentheses. Standard errors are clustered at the subject level.
DG: Dictator Game; RT: Risk Task.
*** p<0.01, ** p<0.05, * p<0.10.



(a) Indicated preferences (b) Beliefs about others' preferences

Figure 4: Preferences for DM under Treatment GA

subjects' preferences in treatment GA should be for the highest type to be appointed as the DM. Figure 4 presents the subjects' preferences for their DM's type under treatment GA (panel a) and their beliefs about the preferences of the other group members (panel b). The majority of the subjects (77.6%) prefer to have the individual who made the highest transfer in the dictator game to be the DM of their group. Moreover, the majority of the subjects (76.1%) believe that the other members of their group prefer to appoint the individual who made the highest transfer as the DM.

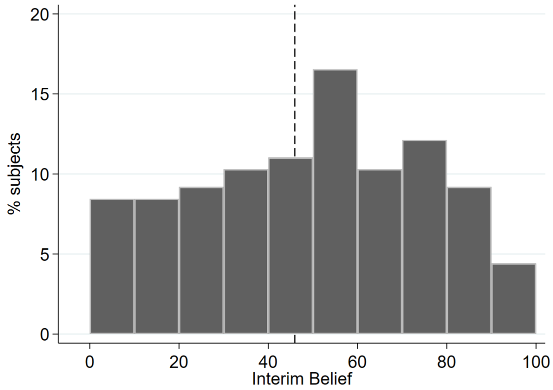
2.3.2 Analysis of interim beliefs

We next examine the members' interim beliefs after they observe the appointment mechanism but prior to observing the DM's outcomes. In all of our analyses, belief is a variable that takes an integer value in $[0, 100]$, where a higher belief implies that the member thinks the DM is more likely to have chosen high effort (e_H). Figure 5 presents the distributions of the members' interim beliefs by treatment. In each panel, the dashed line represents the mean interim belief.

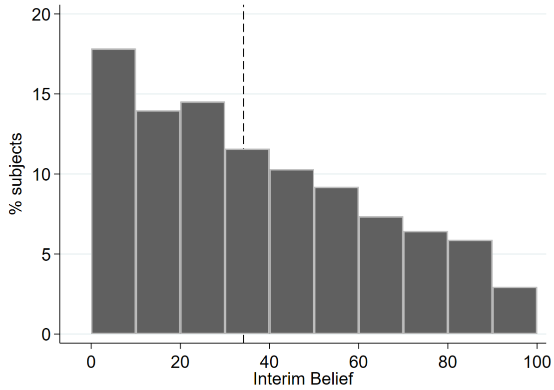
The histograms in Figure 5 suggest that group members respond to the mechanism used to appoint the DM, as stated in Hypothesis 1. In treatment RA, the DM is randomly assigned and the members' beliefs are approximately centered on 50%, with a mean of 45.94% (panel a). In contrast, the distribution of interim beliefs is highly skewed to the right in treatment LA with a mean of 34.15% (panel b), and to the left in treatment HA with a mean of 57.40% (panel c). When the DM is appointed based on the preferences of the group in treatment GA (panel d), the distribution of interim beliefs shifts slightly to the right relative to that in treatment RA, and the average interim belief increases to 48.65% which is lower than that in treatment HA.[24]

Table 3 presents OLS estimates for the regressions of interim beliefs against treatment variables, controlling for Game 1, order effects (in columns 1 and 3), and individual fixed

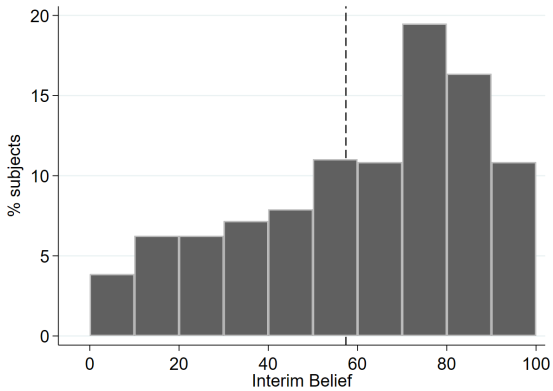
[24]Pairwise comparisons provide initial tests of Hypothesis 1. The Kolmogorov-Smirnov tests reveal that the distributional differences are statistically significant in all comparisons (RA vs. LA: p-value $<$ 0.001; RA vs. HA: p-value $<$ 0.001; GA vs. HA: p-value $<$ 0.001) except one (RA vs. GA: p-value $=$ 0.664). Similarly, Wilcoxon signed-rank tests reject the null hypotheses that the average interim belief is equal between treatments RA and LA (p-value $<$ 0.001), RA and HA (p-value $<$ 0.001), RA and GA (p-value $=$ 0.006) and GA and HA (p-value $<$ 0.001).



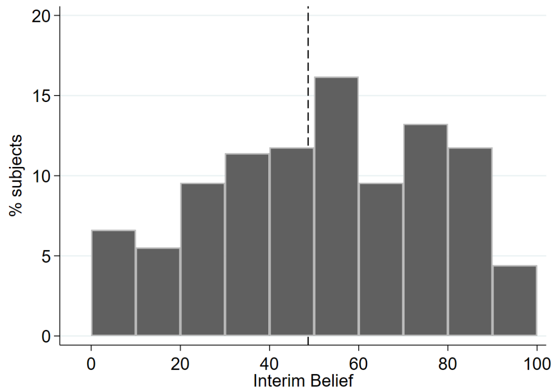
(a) Treatment RA



(b) Treatment LA



(c) Treatment HA



(d) Treatment GA

Figure 5: Distributions of group members' interim beliefs

effects (in columns 2 and 4). In all the specifications, treatment RA is the comparison group. The last row presents the results of a Wald test of equality between treatments GA and HA. The coefficient estimates in columns (1) and (2) support our conclusions from the non-parametric analysis. We also find that the members' beliefs are on average lower in Game 1 treatments than in Game 0 treatments. This difference is statistically significant in both columns (1) and (2) (p-value = 0.005 in both columns).

In columns (3) and (4), we control for the subjects' own decision as a DM. A subject who chooses to exert high effort when placed in the position of the DM under a specific appointment mechanism is also more likely, as a group member, to expect the DM to choose high effort under the same appointment mechanism. This effect is statistically significant (p-value < 0.001 in both columns).[25] The treatment effects remain similar in

[25] Consistent with this finding, the consensus effect is also present in members' beliefs about the DM's behavior in the dictator game. Specifically, in additional regression analyses of members' beliefs about

Table 3: Regression of members' interim belief

	Dependent variable: Interim belief			
Variables	(1)	(2)	(3)	(4)
Treatment LA	-13.268^{***}	-13.268^{***}	-12.237^{***}	-12.646^{***}
	(1.417)	(1.416)	(1.403)	(1.372)
Treatment HA	9.982^{***}	9.982^{***}	8.950^{***}	9.359^{***}
	(1.311)	(1.309)	(1.263)	(1.246)
Treatment GA	2.717^{**}	2.717^{**}	1.857	2.198^{*}
	(1.355)	(1.353)	(1.273)	(1.271)
Chooses high effort as DM			23.382^{***}	14.109^{***}
			(1.848)	(1.588)
% endowment invested in RT	-0.104^{**}		-0.073^{**}	
	(0.043)		(0.036)	
Game 1	-2.952^{***}	-2.952^{***}	-1.405	-2.018^{**}
	(1.041)	(1.040)	(0.955)	(0.964)
Constant	59.182^{***}	48.890^{***}	48.117^{***}	44.066^{***}
	(3.916)	(1.237)	(3.534)	(1.243)
Order Effects	Y	N	Y	N
Individual FE	N	Y	N	Y
Observations	1,632	1,632	1,632	1,632
# subjects (clusters)	272	272	272	272
R-squared	0.137	0.251	0.286	0.305
Test of GA = HA				
test statistic	5.604	5.610	5.738	5.860
p-value	0.000	0.000	0.000	0.000

Robust standard errors clustered at the subject level in parentheses. For all regressions, treatment RA is the reference treatment.
RT: Risk Task.
*** p<0.01, ** p<0.05, * p<0.10.

both direction and magnitude after controlling for the consensus effect, although the estimates for treatment GA are now statistically insignificant in column (3) and marginally statistically significant in column (4) (p-values = 0.159 and 0.093, respectively).

the DM's dictator game behavior, we find a statistically significant positive relationship between the members' own giving behavior and their beliefs about the DM's giving behavior in the dictator game.

We summarize our results in support for Hypothesis 1 as follows:

Result 1: Group members respond to the appointment mechanism in their interim beliefs. The interim beliefs are the lowest in treatment LA and the highest in treatment HA. The interim beliefs in treatment RA are lower than those in treatment GA.

2.3.3 Analysis of posterior beliefs

Estimation strategy for posterior beliefs. To test Hypothesis 2 and analyze updating behavior, we estimate the following equation:

$$\mathrm{logit}(\hat{\phi_i^{\Psi}}|_Q) = \delta\,\mathrm{logit}(\hat{\mu_i^{\Psi}}) + \gamma_G\, I(Q = Q_H)\cdot \mathrm{logit}(p) + \gamma_B\, I(Q = Q_L)\cdot \mathrm{logit}(1-p) + \varepsilon_i, \quad (2)$$

where $\mathrm{logit}(x) \equiv \log(x)/\log(1-x)$ and ε_i captures non-systematic errors. This specification allows us to determine the weights members place on their interim beliefs and the signals they receive.[26] Note that $\delta = \gamma_G = \gamma_B = 1$ corresponds to the Bayesian benchmark. Hypothesis 2 states that $\delta = \gamma_G = \gamma_B = 1$ for each appointment mechanism.

Any deviation in the parameters from 1 is interpreted as non-Bayesian updating behavior. Appendix E provides detailed explanations of the interpretations of these parameters. First, δ captures the weight that group members place on their interim belief in the updating process. If $\delta < 1$ ($\delta > 1$), then members suffer from *base-rate neglect* (*confirmatory bias*) in that they place too little (too much) weight on their interim belief.

Next, γ_G and γ_B capture the extent to which members respond to signals of good outcome and bad outcome from the DM, respectively. $\gamma_G > 1$ or $\gamma_B > 1$ implies that members are, on average, over-responsive to a good or a bad signal, respectively, relative to a Bayesian. Specifically, biased members attribute the corresponding outcome more to the DM's decision as compared to unbiased Bayesian members. On the other hand, $\gamma_G < 1$ or $\gamma_B < 1$ implies that members are conservative in their response to a good or a bad signal, respectively, and attribute the corresponding outcome more to the DM's luck as compared to unbiased Bayesian members.

Finally, we can also capture asymmetric updating of beliefs, i.e., asymmetric attribu-

[26] See, e.g., Grether (1980), Möbius et al. (2014), Ambuehl and Li (2018), Buser et al. (2018), and Coutts (2019) for similar estimation approaches.

tion of outcomes to the DM's decision and luck. If $\gamma_G > \gamma_B$ ($\gamma_G < \gamma_B$), then members are more likely to attribute a good (bad) outcome to the DM's decision and a bad (good) outcome to luck.

Estimating deviations from Bayes' rule. We now estimate equation (2) using ordinary least squares (OLS) to analyze the biases that members suffer from when updating their beliefs.[27] Figure D.1 of Appendix D.3 shows the distribution of subjects who update their beliefs inconsistently (i.e., in the opposite direction to that predicted by Bayes' rule) or not at all. The inclusion of these observations in the analysis may result in biased or incorrect conclusions, particularly if these subjects are reporting beliefs that do not genuinely reflect their true posterior beliefs. Hence, for the remainder of the analysis, we exclude a subject if 25% or more of their posterior beliefs are inconsistent (44 out of 272 subjects in total) or if they report a posterior belief equal to the interim belief across all six rounds of the experiment (23 subjects in total). These two groups jointly constitute 24.6% of the sample. Note that these numbers are largely in line with what is found in the literature (see, e.g., Möbius et al., 2014; Coutts, 2019; Barron, 2020).[28]

Table 4 presents the regression results of members both at the pooled level (column 1) and at the treatment level (columns 2 to 5). As a test of Hypothesis 2, our primary interest is to examine whether the coefficients are different from 1. Hence, asterisks are used in the table to indicate whether a coefficient is statistically significantly different from 1.

Column (1) shows that group members are biased in their belief-updating process.

[27] Note that the logit function is only defined for beliefs in (0,100). Instead of excluding observations of subjects who state 0 or 100 as their interim or posterior belief about the DM, we take the logit of 0.01 or 99.99 as an approximation. Also, one potential concern with estimating (2) using OLS is that the estimates are biased if there are measurement errors in the subjects' reported beliefs. For example, subjects could make mistakes or are imprecise when reporting their beliefs. For robustness, we also consider an alternative specification where the appointment mechanisms are used as instruments for the logit of members' interim beliefs for the analysis at the pooled level. This instrumental-variable (IV) approach requires that the appointment mechanisms are a strong predictor of the members' interim beliefs (as we show in Section 2.3.2) and do not have a separate direct effect on their posterior beliefs. We find that the IV estimates lead to similar conclusions. Details of the results from the IV regression analysis can be found in Appendix D.2.

[28] We present the analyses including these subjects in Table D.5 of Appendix D.3. The main results remain robust despite an attenuation of the coefficient estimates of γ_G and γ_B. Tables D.6 and D.7 of Appendix D.3 also reveal that our results are robust to using different criteria for excluding inconsistent updaters and non-updaters.

Table 4: Regression of members' posterior beliefs

	Dependent variable: Logit(posterior)				
	(1)	(2)	(3)	(4)	(5)
Variables	Pooled	RA	LA	HA	GA
δ : logit(interim belief)	0.695***	0.764***	0.692***	0.703***	0.529***
	(0.039)	(0.071)	(0.054)	(0.058)	(0.135)
γ_G : Good outcome $\times$ logit(p)	0.751***	0.744***	0.622***	0.847*	0.798**
	(0.051)	(0.089)	(0.079)	(0.081)	(0.098)
γ_B : Bad outcome $\times$ logit($1-p$)	0.966	0.932	1.058	0.946	0.876
	(0.067)	(0.092)	(0.117)	(0.072)	(0.114)
Observations	2,460	410	820	820	410
# subjects (clusters)	205	205	205	205	205
R-squared	0.608	0.686	0.651	0.583	0.421
Test of $\gamma_G = \gamma_B$					
test statistic	-3.190	-1.588	-3.065	-1.081	-0.512
p-value	0.002	0.114	0.002	0.281	0.609

Robust standard errors clustered at the subject level in parentheses. This analysis excludes subjects classified as inconsistent or non-updaters.
*** p<0.01, ** p<0.05, * p<0.10. Null hypothesis is coefficient = 1.

The estimate for δ suggests that they suffer from base-rate neglect on average (test of $\delta = 1$: p-value < 0.001).[29] The estimate for γ_G suggests that after controlling for the weight members place on their interim beliefs, members are conservative in their responses to good outcomes. That is, they attribute good outcomes to luck more than a Bayesian would and this effect is statistically significant (test of $\gamma_G = 1$: p-value $<$ 0.001). However, there is no statistically significant evidence that members respond to bad outcomes differently from a Bayesian (test of $\gamma_B = 1$: p-value $= 0.608$). Hence, relative to the Bayesian benchmark, group members give too little credit for the DM's success but the right amount of blame for the DM's failure.

The last two rows of Table 4 present the results of a Wald test of equality between γ_G and γ_B, giving us a test of the presence of an asymmetric attribution bias. Overall, members update their beliefs about the DM asymmetrically (i.e., $\gamma_G < \gamma_B$). They tend to attribute good outcomes more to luck and, relatively, bad outcomes more to the DM's

[29] An alternative and less restrictive specification to (2) is to allow δ (the weight placed on interim beliefs) to vary with the signal received by the members. However, estimating δ_G and δ_B separately, we find that both of these parameter estimates are statistically significantly different from 1 (p-values < 0.001 for both), but the difference between them is not statistically significant (p-value $= 0.692$). Moreover, this less restrictive specification does not affect the estimates of γ_G and γ_B.

decision. This effect is statistically significant (p-value = 0.002).[30]

We next analyze the members' updating behavior across the different appointment mechanisms. The coefficient estimates in columns (2)-(5) of Table 4 reveal that biases similar to the ones observed at the pooled level exist at the treatment level. Under each appointment mechanism, members consistently suffer from base-rate neglect, attribute good outcomes more to luck, and treat bad outcomes no differently from a Bayesian.[31,32] The asymmetry observed in the attribution of outcomes is statistically significant in treatment LA only (Wald tests of $\gamma_G = \gamma_B$: p-values = 0.002, 0.114, 0.281 and 0.609, respectively, for treatments LA, RA, HA and GA).[33]

In summary, we do not find support for Hypothesis 2. Members are not Bayesian when updating their beliefs after observing their DM's outcomes.

Result 2: On average, group members exhibit similar biases in their updating behavior under all the appointment mechanisms. They suffer from base-rate neglect in their updating behavior. Compared to the Bayesian benchmark, members attribute good outcomes more to luck, but their average response to bad outcomes is not different from Bayesian.

We next consider the relationship between subjects' effort choices (which are determined by their types) and their updating behavior. The results in Section 2.3.2 reveal that subjects' effort choices are correlated with their interim beliefs. Our aim here is to test whether subjects who exert high effort as DMs update their beliefs differently to those who exert low effort after controlling for their interim beliefs.

Table 5 reports separate parameter estimates of (2) based on whether the subjects

[30]We also consider heterogeneity in updating behavior using finite mixture model analyses in Appendix D.4. The results show that the majority of belief updates in the sample is characterized by base-rate neglect, under-responsiveness to DMs' outcomes, and an asymmetric attribution of DMs' outcomes to decision and luck.

[31]Note that base-rate neglect and under-inference are commonly observed in other studies adopting a similar estimation strategy to our paper (Benjamin, 2019). However, we do not observe under-inference relative to a Bayesian in response to bad outcomes.

[32]Comparing the magnitudes of the biases across the appointment mechanisms, we fail to reject the null hypotheses that the estimates for δ, γ_G, and γ_B are jointly equal to one another (Wald tests of $\delta^{RA} = \delta^{LA} = \delta^{HA} = \delta^{GA}$: p-value = 0.395; $\gamma_G^{RA} = \gamma_G^{LA} = \gamma_G^{HA} = \gamma_G^{GA}$: p-value = 0.110; and $\gamma_B^{RA} = \gamma_B^{LA} = \gamma_B^{HA} = \gamma_B^{GA}$: p-value = 0.686).

[33]Note that the magnitude of the asymmetry is the largest in treatment LA, which is driven both by the estimate for γ_G being the lowest and that for γ_B being the highest in this treatment. This suggests that across the four treatments, group members are the least likely to believe that good outcomes result from a choice of high effort in the case when the most selfish individual is appointed to be the DM.

Table 5: Regression of members' posterior beliefs based on effort choice as DMs

	Dependent variable: Logit(posterior)	
	(1)	(2)
Variables	Chose low effort	Chose high effort
δ : logit(interim belief)	0.710***	0.603***
	(0.048)	(0.073)
γ_G : Good outcome $\times$ logit(p)	0.698***	0.957
	(0.059)	(0.110)
γ_B : Bad outcome $\times$ logit($1-p$)	0.923	0.951
	(0.080)	(0.100)
Observations	1,646	814
# subjects (clusters)	190	125
R-squared	0.626	0.553
Test of $\gamma_G = \gamma_B$		
t-statistic	-2.568	0.049
p-value	0.011	0.961

Robust standard errors clustered at the subject level in parentheses. This analysis excludes subjects classified as inconsistent or non-updaters.
*** p<0.01, ** p<0.05, * p<0.10. Null hypothesis is coefficient = 1.

have chosen low effort (column 1) or high effort (column 2) as DMs within a given round in the investment task. The estimates of δ and γ_B are not statistically significantly different between columns (1) and (2) (p-values = 0.222 and 0.818, respectively). However, the estimate of γ_G is statistically significantly different between the two columns (p-value = 0.035). While the estimate for γ_G is statistically significantly less than 1 in column (1) (p-value $<$ 0.001), it is not different from 1 in column (2) (p-value = 0.697). Hence, regardless of their effort choices as DMs in a given round of the task, subjects suffer from base-rate neglect ($\delta < 1$) and are no different from a Bayesian in their response to bad outcomes ($\gamma_B = 1$) on average. However, in a given round of the investment task, those individuals who choose low effort as DMs are more likely to attribute good outcomes to luck when they make decisions as group members.[34] This suggests that the consensus

[34] Our results here complement those of Di Tella et al. (2015) who show that individuals avoid altruistic actions by distorting their beliefs about others' altruism.

effect that we find to be driving members' interim beliefs also appears to be shaping their updating behavior.

3 Experiment 2

3.1 Motivation and Design

The results from Experiment 1 suggest that subjects use their own behavior as the basis for forming and updating their beliefs about others. We design Experiment 2 to investigate this further. In treatment S (single role), subjects are informed at the beginning of the experiment whether they have been assigned as the DM or as a group member. They make incentivized decisions only in their assigned roles. Specifically, they are asked to make their investment decisions (only if they are the DM) or to report their beliefs about the DM's decisions (only if they are the member). After reporting their beliefs at the end of each round, members are asked to indicate, hypothetically, what their investment decision would have been if they were the DM. This structure allows us to examine the relationship between individuals' effort choices as DMs and their beliefs as members, while mitigating the impact that experience may have on their beliefs.

The key question we ask in Experiment 2 is whether we continue to observe the consensus effect in members' attribution of outcomes. Given the well-documented evidence on the consensus effect in driving behavior (e.g., Ross et al., 1977; Marks and Miller, 1987; Engelmann and Strobel, 2000, 2012), we expect the relationship between subjects' behavior and their attribution of outcomes to persist even if they do not make choices as decision makers prior to stating their beliefs.[35] Consequently, we expect to observe similar biases in the overall attribution of outcomes when decisions are no longer elicited using the strategy method.[36]

The sessions were conducted online during a lockdown amid the COVID-19 pan-

[35]Even if we do not elicit incentivized effort choices from members, a consensus effect may still exist if members consider in their minds what their actions would have been as DMs before stating their beliefs.

[36]Brandts and Charness (2011) show that the strategy method typically finds qualitatively similar effects compared to the direct-response method.

demic.[37] Subjects were recruited from a similar subject pool at the University of Melbourne, with between 18 and 24 subjects in each virtual session.[38] The experiments were programmed using oTree (Chen et al., 2016). For each session, all the subjects were admitted into a Zoom meeting with their videos and microphones turned off. They were provided with separate links for the instructions to each part of the experiment, and the experimenter read out the instructions in the Zoom meeting. A total of 297 subjects (99 DMs and 198 group members) participated in treatment S, and each session lasted about 60 minutes on average.

The implementation of treatment S was different on two other dimensions to accommodate the different and shorter nature of the online sessions. First, since we do not find any significant evidence in Experiment 1 that members' updating behavior depends on the appointment mechanism (Result 2), we removed this treatment variation by using only the random-appointment mechanism (RA) to determine the DM of each group. Second, each group participated in only three rounds of the investment task (instead of six), remaining in the same group and role for all three rounds. In two of the rounds, we used the same parameters as in Game 0 and Game 1 from Experiment 1. We introduced Game 2 as the third set of parameters. Relative to Game 1, both investments in Game 2 provide a higher return of 900 ECU if they succeed and the same low return of 150 ECU if they fail.[39] We randomized the order of these parameters across the groups within each session.

Given the differences in format between Experiment 1 and Experiment 2, we also ran treatment D (dual role) online to have a comparable benchmark against which to evaluate subjects' behavior in treatment S. In this treatment, subjects made decisions both in the roles of the DM and a group member (as in Experiment 1). 206 subjects participated in treatment D.

[37]The same experimenter ran both Experiment 1 and Experiment 2. Instructions for Experiment 2 are available in Appendix A.2. Experiment 2 (the design, treatments, empirical strategy, and power calculations) is pre-registered on the AEA RCT Registry (AEARCTR-0006519: https://doi.org/10.1257/rct.6519-1.1).

[38]Unlike Experiment 1, both domestic and international students were recruited for the sessions in Experiment 2.

[39]Our theoretical prediction is that DMs are more likely to choose high effort in Game 2 as compared to in Game 1.

3.2 Results

Table D.9 in Appendix D.5 reveals that there are significant differences in the subject pools between Experiment 1 and Experiment 2. Subjects in Experiment 2 are on average slightly older, less likely to be Australian or majoring in economics, more likely to be a postgraduate student, and more experienced with economics experiments. However, with respect to the behavioral variables, the subjects do not differ in their decisions in the dictator game or risk task between Experiment 1 and Experiment 2.

We are mainly interested in analyzing members' posterior beliefs in treatment S, to see whether the same type of biases as in Experiment 1 exist and whether we find evidence of a consensus effect.[40] To test for the consensus effect, we use the hypothetical choices that members make when they are asked what their investment decision would have been if they were the DM. Table D.10 in Appendix D.5 reveals that there is a statistically significant and positive relationship between members' hypothetical choices as DMs and their incentivized decisions in the dictator game. Hence, the hypothetical answers given by the members seem to be indicative of their underlying preferences.

We start by analyzing the posterior beliefs elicited in Round 1 of treatment S, which provides the cleanest examination of members' biases absent any experience in the decision-making process as DMs.[41] Columns (1) and (2) of Table 6 present parameter estimates of equation (2) based on whether members indicate, hypothetically, that they would have chosen low effort or high effort, respectively, if they were the DM of the group. Columns (3) and (4) report parameter estimates including members' belief updates in all rounds of the investment task. In the table, we also present p-values of pairwise comparisons of parameter estimates between columns (1) and (2), and between columns (3) and (4).

Columns (1) and (2) reveal that members who would have chosen low effort as DMs

[40] Note that in Experiment 2, we consider members' unconditional beliefs as *prior* beliefs since, unlike Experiment 1, there are no variations in the appointment mechanisms that would have otherwise influenced the members' unconditional beliefs.

[41] While we only report results on members' updating behavior here, we also examine the DMs' effort choices and members' prior beliefs in Experiment 2. Our results are consistent with our findings from Experiment 1. Specifically, there exists a positive relationship between the DM's decision to choose high effort in the investment task and their giving behavior in the dictator game. Moreover, using members' hypothetical choices as DMs, we continue to find evidence of a consensus effect. On average, members who would have chosen high effort as DMs have higher prior beliefs about the DM choosing high effort.

Table 6: Regression of members' posterior beliefs based on hypothetical effort choice as DMs (treatment S)

	Dependent variable: Logit(posterior)					
	Round 1 only			All rounds		
	(1)	(2)	(1) vs. (2)	(3)	(4)	(3) vs. (4)
Variables	Low effort	High effort	p-value	Low effort	High effort	p-value
δ : logit(prior belief)	0.560***	0.395***	0.191	0.483***	0.607***	0.254
	(0.110)	(0.062)		(0.079)	(0.077)	
γ_G : Good outcome $\times$ logit(p)	0.490*	0.995	0.092*	0.601*	1.068	0.085*
	(0.262)	(0.143)		(0.210)	(0.173)	
γ_B : Bad outcome $\times$ logit($1-p$)	0.774	0.636*	0.688	1.181	0.759	0.082*
	(0.288)	(0.192)		(0.188)	(0.171)	
Observations	204	92		656	232	
# subjects (clusters)	102	46		137	71	
R-squared	0.370	0.532		0.392	0.488	

Robust standard errors clustered at the subject level in parentheses. This analysis excludes subjects classified as inconsistent or non-updaters. All columns control for members' hypothetical effort choices in treatment S. Columns (1) and (2) restrict the analysis to the first round of updates only.
*** p<0.01, ** p<0.05, * p<0.10. Null hypothesis is coefficient = 1.

are more likely to attribute the DM's good outcomes to luck relative to a Bayesian (p-value = 0.054), while those who would have chosen high effort are no different from a Bayesian in their attribution of good outcomes (p-value = 0.435). Overall, members who would have chosen low effort are more likely to attribute good outcomes to luck than those who would have chosen high effort (p-value = 0.092).[42] These findings are consistent with those from Experiment 1 (Table 5). Interestingly, we now find that members who would have chosen high effort as DMs are more likely to attribute bad outcomes to luck relative to a Bayesian (p-value = 0.064). However, there is no statistically significant difference in the attribution of bad outcomes between members who would have chosen high effort and those who would have chosen low effort (p-value = 0.688).

Columns (3) and (4) of Table 6 reveal that, when we consider all rounds of the task, members who would have chosen low effort as DMs attribute good outcomes more to luck both relative to a Bayesian (p-value = 0.060) and relative to those who would have chosen high effort as DMs (p-value = 0.085). These results are consistent with those observed in columns (1) and (2), as well as Experiment 1. Moreover, we now find that members who would have chosen high effort as DMs are more likely to attribute bad outcomes to luck

[42] We show later that the results in Experiment 2 have similar patterns as those in Experiment 1. The noisier behaviour of subjects in the virtual environment, however, may have led to some estimates being less precise in Experiment 2.

as compared to members who would have chosen low effort as DMs (p-value = 0.082).[43]

Finally, we compare members' updating behavior across treatment S, treatment D, and Experiment 1. Tables D.11 and D.12 in Appendix D.5 present comparisons of members' updating behavior between treatments S and D, separately by members' effort choices and at the pooled level, respectively. The last column of Table D.12 also provides p-values from tests of differences in parameter estimates between Experiment 1 and Experiment 2. The tables provide two main insights. First, we do not find any systematic differences between treatments S and D in members' attribution of outcomes, both when we compare subjects separately based on their effort choices as DMs, and at the pooled level. Second, despite the differences in the experimental design, format, and subject pool, we do not find any statistically significant differences in members' attribution of outcomes between Experiment 1 and Experiment 2.

Hence, we conclude that there is no evidence to suggest that members' biases in their attribution of the DM's outcomes are driven by whether or not they have experience in the decision-making process as DMs. With and without the experience of acting as a decision maker, the same attribution biases exist and seem to be driven by a consensus effect.

4 Conclusion

In many environments, the determinants of outcomes are not observable. What beliefs do individuals hold in such circumstances about the determinants of others' outcomes? Do they attribute the outcomes to luck or to the decisions made? Do the beliefs depend on the outcome, i.e., whether the outcome is good or bad? These are the questions we address in this paper.

Our results reveal that members suffer from biases in the way they attribute outcomes to luck versus the choices made. Moreover, members treat good and bad outcomes differently in the sense that while they attribute good outcomes more to luck as compared

[43]Note that this result is consistent with a consensus effect being present also for those members who would have chosen high effort as DMs.

to a Bayesian, their response to bad outcomes is no different from a Bayesian. This asymmetry implies that the credit decision makers receive for good outcomes is less than the blame they get for bad outcomes.

In Experiment 1, we find that group members exhibit similar biases under all the mechanisms used to appoint the decision maker. However, we find that biases in updating behavior tend to be driven by those subjects who choose low effort as decision makers. Interestingly, the consensus effect we detect affects both initial beliefs and updating behavior. In Experiment 2, we show that the same type of biases exist even if the subjects do not have experience as decision makers.

Determining the systematic biases that individuals may have in the way they process new information and update their beliefs about the decisions of others is critical in a wide range of economic and social interactions. One general implication of our study is that the biases we identify may affect the generosity of decision makers in environments where social preferences matter. For example, they may act less generously if they know that they will not receive sufficient credit for good outcomes. The biases may also affect decision makers' willingness to take risk. For instance, if business or political leaders are aware that they are given relatively more blame for their failures than credit for their successes, then this may perpetuate a culture of failure avoidance. Such a 'fear of failure' culture may reduce their incentives to exert costly effort or their tolerance towards risk.

Our study identifies the biases which exist in the evaluation of others' decisions specifically in contexts where prosocial preferences play a key role in decision making. In future research, it would be interesting to understand whether the same type of biases exist in other contexts. For example, if performance in a skill-based task is important for leadership, do we observe that the same type of biases emerge in evaluation? Or, if we remove the anonymity of the decision maker, to what extent do in-group versus out-group considerations affect leadership evaluation? Answering these questions would broaden our understanding of biases in performance evaluation.

References

Alesina, A., G.-M. Angeletos (2005). Fairness and redistribution. *American Economic Review*, 95(4):960–980.

Ambuehl, S., S. Li (2018). Belief updating and the demand for information. *Games and Economic Behavior*, 109:21–39.

Ames, D. R. (2004). Inside the mind reader's tool kit: Projection and stereotyping in mental state inference. *Journal of Personality and Social Psychology*, 87(3):340–353.

Barron, K. (2020). Belief updating: Does the 'good-news, bad-news' asymmetry extend to purely financial domains? *Experimental Economics*:1–28.

Bénabou, R., J. Tirole (2010). Individual and corporate social responsibility. *Economica*, 77(305):1–19.

Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. In B. D. Bernheim, S. DellaVigna, D. Laibson (eds.), *Handbook of Behavioral Economics: Applications and Foundations*, volume 2, chapter 2. North-Holland: Elsevier, pp. 69–186.

Bertrand, M., S. Mullainathan (2001). Are CEOS rewarded for luck? The ones without principals are. *The Quarterly Journal of Economics*, 116(3):901–932.

Brandts, J., G. Charness (2011). The strategy versus the direct-response method: A first survey of experimental comparisons. *Experimental Economics*, 14(3):375–398.

Brock, J. M., A. Lange, E. Y. Ozbay (2013). Dictating the risk: Experimental evidence on giving in risky environments. *American Economic Review*, 103(1):415–37.

Brownback, A., M. A. Kuhn (2019). Understanding outcome bias. *Games and Economic Behavior*, 117:342–360.

Buser, T., L. Gerhards, J. van der Weele (2018). Responsiveness to feedback as a personal trait. *Journal of Risk and Uncertainty*, 56(2):1–28.

Cappelen, A. W., J. Konow, E. Ø. Sørensen, B. Tungodden (2013). Just luck: An experimental study of risk taking and fairness. *American Economic Review*, 103(4):1398–1413.

Charness, G., D. I. Levine (2007). Intention and stochastic outcomes: An experimental study. *The Economic Journal*, 117(522):1051–1072.

Chen, D. L., M. Schonger, C. Wickens (2016). oTree–An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Cole, S., A. Healy, E. Werker (2012). Do voters demand responsive governments? Evidence from Indian disaster relief. *Journal of Development Economics*, 97(2):167–181.

Coutts, A. (2019). Good news and bad news are still news: Experimental evidence on belief updating. *Experimental Economics*, 22(2):369–395.

Darley, J. M., P. H. Gross (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, 44(1):20–33.

Dawes, R. M. (1989). Statistical criteria for establishing a truly false consensus effect. *Journal of Experimental Social Psychology*, 25(1):1–17.

Deffains, B., R. Espinosa, C. Thöni (2016). Political self-serving bias and redistribution. *Journal of Public Economics*, 134:67–74.

Di Tella, R., R. Perez-Truglia, A. Babino, M. Sigman (2015). Conveniently upset: Avoiding altruism by distorting beliefs about others' altruism. *American Economic Review*, 105(11):3416–42.

Edelson, M. G., R. Polania, C. C. Ruff, E. Fehr, T. A. Hare (2018). Computational and neurobiological foundations of leadership decisions. *Science*, 361(6401).

Eil, D., J. M. Rao (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–138.

Engelmann, D., M. Strobel (2000). The false consensus effect disappears if representative information and monetary incentives are given. *Experimental Economics*, 3(3):241–260.

Engelmann, D., M. Strobel (2012). Deconstruction and reconstruction of an anomaly. *Games and Economic Behavior*, 76(2):678–689.

Erkal, N., L. Gangadharan, B. H. Koh (2020). Replication: Belief elicitation with quadratic and binarized scoring rules. *Journal of Economic Psychology*, 81.

Ertac, S. (2011). Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback. *Journal of Economic Behavior and Organization*, 80(3):532–545.

Ertac, S., M. Y. Gurdal (2012). Deciding to decide: Gender, leadership and risk-taking in groups. *Journal of Economic Behavior and Organization*, 83(1):24–30.

Exley, C. L. (2015). Excusing selfishness in charitable giving: The role of risk. *The Review of Economic Studies*, 83(2):587–628.

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

Galeotti, F., D. J. Zizzo (2018). Identifying voter preferences: The trade-off between honesty and competence. *European Economic Review*, 105:27–50.

Galinsky, A. D., G. B. Moskowitz (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, 78(4):708–724.

Garrett, N., T. Sharot (2017). Optimistic update bias holds firm: Three tests of robustness following Shah et al. *Consciousness and Cognition*, 50:12–22.

Gauriot, R., L. Page (2019). Fooled by performance randomness: Overrewarding luck. *Review of Economics and Statistics*, 101(4):658–666.

Gigerenzer, G., U. Hoffrage (1995). How to improve bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102(4):684–704.

Gneezy, U., J. Potters (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics*, 112(2):631–645.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125.

Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *The Quarterly Journal of Economics*, 95(3):537–557.

Grossman, Z., D. Owens (2012). An unlucky feeling: Overconfidence and noisy feedback. *Journal of Economic Behavior and Organization*, 84(2):510–524.

Gurdal, M. Y., J. B. Miller, A. Rustichini (2013). Why blame? *Journal of Political Economy*, 121(6):1205–1247.

Haggag, K., D. G. Pope, K. B. Bryant-Lees, M. W. Bos (2019). Attribution bias in consumer choice. *The Review of Economic Studies*, 86(5):2136–2183.

Holmes, D. S. (1968). Dimensions of projection. *Psychological Bulletin*, 69(4):248–268.

Hossain, T., R. Okui (2013). The binarized scoring rule. *The Review of Economic Studies*, 80(3):984–1001.

Kahneman, D., A. Tversky (1973). On the psychology of prediction. *Psychological Review*, 80(4):237–251.

Lange, A., J. Schmitz, C. Schwirplies (2019). Changing perspectives: Local and global cooperation in presence of income inequalities. *Working Paper*.

Leone, A. J., J. S. Wu, J. L. Zimmerman (2006). Asymmetric sensitivity of CEO cash compensation to stock returns. *Journal of Accounting and Economics*, 42:167–192.

Lepsinger, R., A. D. Lucia (2009). *The art and science of 360 degree feedback*. John Wiley & Sons.

Lord, C. G., L. Ross, M. R. Lepper (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37(11):2098–2109.

Marks, G., N. Miller (1987). Ten years of research on the false-consensus effect: An empirical and theoretical review. *Psychological Bulletin*, 102(1):72–90.

Miller, D. T., M. Ross (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, 82(2):213–225.

Möbius, M. M., M. Niederle, P. Niehaus, T. S. Rosenblat (2014). Managing self-confidence. *Working Paper*.

Nisbett, R. E., E. Borgida (1975). Attribution and the psychology of prediction. *Journal of Personality and Social Psychology*, 32(5):932–943.

Nyhan, B., J. Reifler (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2):303–330.

Palfrey, T. R., S. W. Wang (2012). Speculative overpricing in asset markets with information flows. *Econometrica*, 80(5):1937–1976.

Pettigrew, T. F. (1979). The ultimate attribution error: Extending Allport's cognitive analysis of prejudice. *Personality and Social Psychology Bulletin*, 5(4):461–476.

Plous, S. (1991). Biases in the assimilation of technological breakdowns: Do accidents make us safer? *Journal of Applied Social Psychology*, 21(13):1058–1082.

Price, P. C. (1998). Effects of a relative-frequency elicitation question on likelihood judgment accuracy: The case of external correspondence. *Organizational Behavior and Human Decision Processes*, 76(3):277–297.

Rabin, M., J. L. Schrag (1999). First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics*, 114(1):37–82.

Rey-Biel, P., R. Sheremeta, N. Uler (2018). When income depends on performance and luck: The effects of culture and information on giving. In R. M. Isaac, D. A. Norton (eds.), *Experimental Economics and Culture*, volume 20 of *Research in Experimental Economics*.

Ross, L., D. Greene, P. House (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13(3):279–301.

Rotemberg, J. J. (2014). Models of caring, or acting as if one cared, about the welfare of others. *Annual Review of Economics*, 6:129–154.

Sarsons, H. (2019). Interpreting signals in the labor market: Evidence from medical referrals. *Working Paper*.

Schlag, K. H., J. Tremewan, J. J. van der Weele (2015). A penny for your thoughts: A survey of methods for eliciting beliefs. *Experimental Economics*, 18(3):457–490.

Sharot, T., C. W. Korn, R. J. Dolan (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, 14(11):1475–1479.

Sunstein, C. R., S. Bobadilla-Suarez, S. C. Lazzaro, T. Sharot (2017). How people update beliefs about climate change: Good news and bad news. *Cornell Law Review*, 102(6):1431–1443.

Todd, A. R., G. V. Bodenhausen, J. A. Richeson, A. D. Galinsky (2011). Perspective taking combats automatic expressions of racial bias. *Journal of Personality and Social Psychology*, 100(6):1027–1042.

Tversky, A., D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131.

Tversky, A., D. Kahneman (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, A. Tversky (eds.), *Judgment under uncertainty: Heuristics and biases*, chapter 10. New York: Cambridge University Press, pp. 153–160.

Wolfers, J. (2007). Are voters rational? Evidence from gubernatorial elections. *Working Paper*.

[FOR ONLINE PUBLICATION]

A Experimental instructions

A.1 Experiment 1 (Lab)

Overview of Experiment

Thank you for agreeing to take part in this study which is funded by the Australian Research Council. Please read the following instructions carefully. A clear understanding of the instructions will help you make better decisions and increase your earnings from the experiment.

You will participate in two experiments today: Experiment 1 and Experiment 2.[1] You will receive detailed instructions for each experiment before you participate in them. Note that your decisions in Experiment 2 will not change the earnings that you receive from Experiment 1. You will be informed of the outcomes of both experiments at the end of today's session.

You will be paid for the decisions you make in either Experiment 1 or Experiment 2. This implies that you should carefully consider all of the decisions you make in both experiments as they may determine your earnings. Whether you will be paid for Experiment 1 or Experiment 2 will be randomly determined at the end of the session. Your final payment today will also include a $10 participation fee.

During the experiments, we will be using Experimental Currency Units (ECU). At the end of the session, we will convert the amount you earn into Australian Dollars (AUD) using the following conversion rate: 10 ECU = 1 AUD.

At the end of Experiment 2, you will be asked to fill out a brief questionnaire asking you some general questions. All of the decisions you make in today's session will remain anonymous.

Please do not talk to one another during the experiment. If you have any questions, please raise your hand and we will come over to answer your questions privately.

[1] Note that in the instructions, we use Experiment 1 and Experiment 2 to refer to the two tasks (dictator game and investment task) that subjects participate in.

Experiment 1

You will participate in Experiment 1 in groups of two. The computer will randomly match you with one other person in the room. You will never learn the identity of your partner.

Each of you is given an endowment of 300 ECU, and you are asked to divide this amount between yourself and the person you are matched with.

At the end of today's session, if this experiment is picked for payment, then you will be paid either according to your decision or according to the decision made by your randomly matched partner. The computer will randomly determine whose allocation decision will be implemented.

Example. Suppose you choose to divide your endowment by keeping 200 ECU for yourself and giving 100 ECU to your matched partner. Your matched partner decides to keep 130 ECU and give 170 ECU to you. If, at the end of the experiment, the computer randomly determines that it is the allocation of your matched partner that gets implemented, then your payment will be 170 ECU and your matched partner's payment will be 130 ECU.

Are there any questions? If not, we will proceed with Experiment 1.

Experiment 2

Experiment 2 consists of six identical rounds. At the end of the experiment, if you are paid for Experiment 2, then the computer will randomly pick one of the six rounds for payment.

You will participate in each round in groups of three. At the beginning of each round, the computer will randomly match you with two other people in this room with whom you have not been matched before. You will never learn the identity of your partners. Each round consists of three stages.

Stage 1: Appointment of a group leader.

In this stage, one group member will be assigned to be the leader of the group. There will be four possible methods to determine who is assigned the role of the leader. At the beginning of each round, the computer will reveal which method will be used to determine the leader for that round.

Method 1: One group member will be randomly assigned by the computer to be the leader. Hence, each group member has an equal chance of being assigned the role of the leader.

Method 2: The group member who transferred the lowest amount to his/her matched partner in Experiment 1 will be assigned to be the leader (ties will be broken randomly).

Method 3: The group member who transferred the highest amount to his/her matched partner in Experiment 1 will be assigned to be the leader (ties will be broken randomly).

Method 4: Each individual within the group will be asked to indicate whether you prefer your leader to be someone who has transferred the highest or the lowest amount to his/her matched partner in Experiment 1. The computer will then randomly pick one of the decisions of the group members to implement. If your decision is implemented, then one of your other two group members will be appointed to be the leader based on your preference. Hence, you will not be appointed to be the leader if your decision is implemented.

Example 1. Suppose the leader is appointed using Method 4. In Experiment 1, Player 1 chose to transfer 100 ECU to his/her matched partner, and Player 2 chose to transfer 160 ECU to his/her matched partner. Player 3 indicates that his/her preferred leader is someone who has transferred the lowest amount to his/her matched partner in Experiment 1. If the computer randomly determines that Player 3's decision will be implemented, then Player 1 will be assigned the role of the leader.

You will only need to indicate your preferred leader for Method 4 once, at the beginning of Experiment 2. The same decision will be used whenever Method 4 is being used to determine the appointment of the group leader.

Stage 2: Investment decision by the group leader.

The leader will be given an endowment of 300 ECU. S/he will be asked to choose between two investment options that will affect the payoffs of all group members. Each investment can either fail or succeed. The two investment options have different chances of success/failure. They also have different costs to the leader.

Specifically, the two investments are:

Investment X: This investment costs 250 ECU to the leader. It will succeed with a 75% chance, and fail with a 25% chance.

Investment Y: This investment costs 50 ECU to the leader. It will succeed with a 25% chance, and fail with a 75% chance.

The payoffs to the leader and each group member in this stage of Experiment 2 are calculated as follows:

1. Payoff to leader = 300 ECU − Cost of investment + Returns on investment
2. Payoff to each group member = Returns on investment

Note that the amount that you receive from each investment may be different in each round, and this may affect the final payoffs to the leader and each group member. However, you will always receive a higher payoff if the investment succeeds, and a lower payoff if it fails. Please pay attention to these numbers on the screen in each round.

Figure A.1 shows an example where the returns of each investment options are 200 ECU if the investment succeeds, and 0 ECU if the investment fails, i.e., as shown by the numbers in **red**.

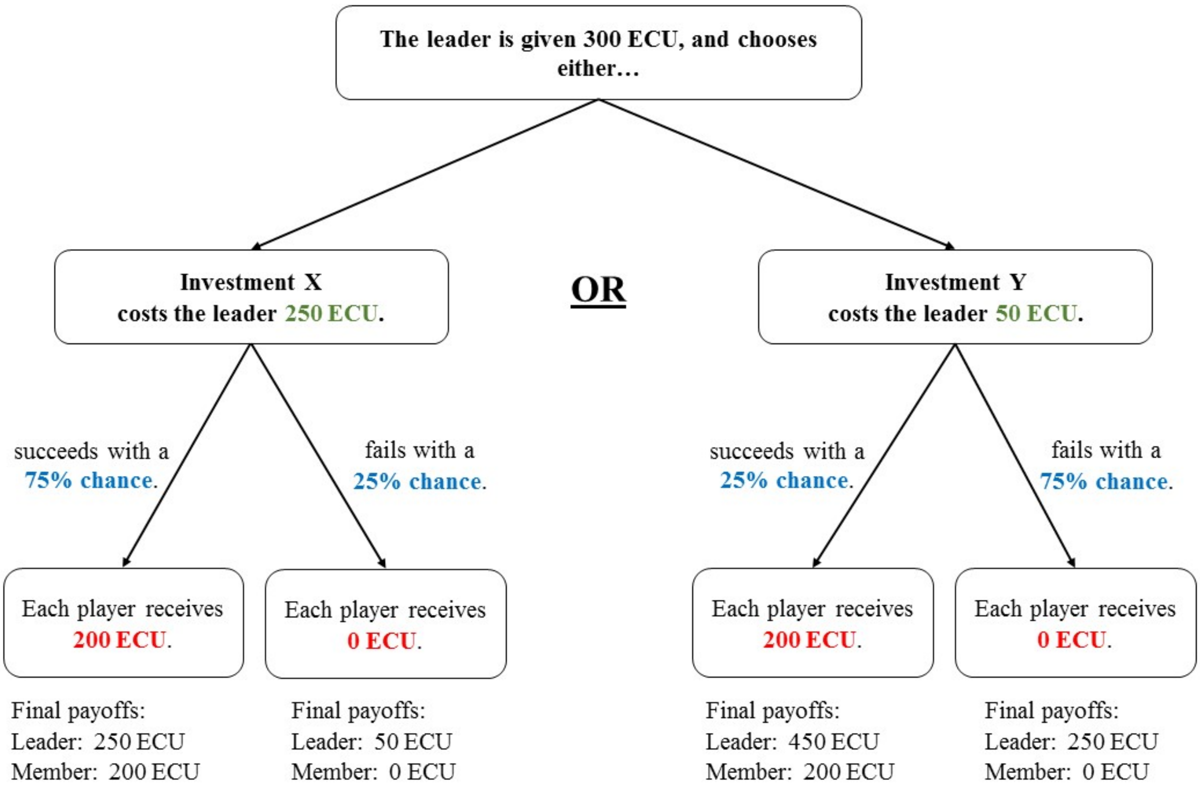


Figure A.1: Investment Options (Example of a Round)

Example 2. Suppose in the round depicted in Figure A.1, the leader chooses Investment X for the group. Then, the investment costs the leader 250 ECU, and will succeed with a 75% chance and fail with a 25% chance. At the end of the experiment, if the investment succeeds, then each group member will receive 200 ECU, and the leader will receive $(300 - 250 + 200) = 250$ ECU for this stage of Experiment 2.

Example 3. Suppose in the round depicted in Figure A.1, the leader chooses Investment Y for the group. Then, the investment costs the leader 50 ECU, and will succeed with a 25% chance and fail with a 75% chance. At the end of the experiment, if the investment fails, then each group member will receive 0 ECU, and the leader will receive $(300 - 50 + 0) = 250$ ECU for this stage of Experiment 2.

You will be informed whether you have been assigned the role of the leader at the end of the experiment. Hence, you will be asked to make an investment decision in Stage 2 of each round assuming that you have been assigned the role of the leader. Your decision will be implemented if you have been assigned the role of the leader for that round.

At the end of the experiment, all group members will learn how much they have received from the chosen investment, but they will not learn the investment decision of the leader.

Stage 3: Beliefs of the other group members.

After you have made your investment decision, you will be asked to predict which investment your leader has chosen, assuming that someone else in your group has been assigned the role of the leader.

Specifically, we would like to know how likely it is in your opinion that the leader has chosen Investment X. *Suppose there were 100 people in the position the leader is in now. How many of them do you think would choose Investment X?*

You will need to choose a number between 0 and 100. A higher number means that you think the leader is more likely to have chosen Investment X.

The specific questions you will be asked are listed below.

Question 1

Suppose there were 100 people in the position the leader is in now. How many of them do you think would choose Investment X?

In Question 2, you are given additional information. You are asked to evaluate the same question with this additional information. Specifically, you should consider whether your guess of the leader's decision will be different, given that you know the outcome of the investment chosen by your leader.

Question 2

> **Suppose you are informed that the investment chosen by your leader has succeeded, and you have therefore received the high payoff.**
>
> Now consider whether your guess will be higher than, lower than, or the same as the one you stated in Question 1. That is, suppose there were 100 people in the position the leader is in now. Given an outcome of high payoff, how many of them do you think have chosen Investment X?
>
> **Suppose you are informed that the investment chosen by your leader has failed, and you have therefore received the low payoff.**
>
> Now consider whether your guess will be higher than, lower than, or the same as the one you stated in Question 1. That is, suppose there were 100 people in the position the leader is in now. Given an outcome of low payoff, how many of them do you think have chosen Investment X?

The computer will randomly select one of these two questions and you will be paid for your response to this question. If Question 2 is chosen for payment, then you will be paid for your answer to the scenario that corresponds to the actual outcome of the investment chosen by your leader.

The section below describes how your payoff in Stage 3 will be determined. This procedure has been used in many other studies. We explain the procedure in detail, but what is most important is that <u>this payoff structure is designed such that it is in your best interest to report your true belief about your leader's decision</u>.

Your payment for the question randomly chosen by the computer is determined as follows. You will receive 10 ECU with some chance. Your chance of receiving 10 ECU depends on your answer and the leader's decision. The closer your guess is to the actual decision made by your leader, the higher is your chance of receiving the fixed payment of 10 ECU.

Specifically, your chance of receiving 10 ECU is determined by the following formula:

$$\text{Chance of receiving 10 ECU} = \left[1 - \left(\frac{x - \text{your guess}}{100}\right)^2\right] \times 100.$$

x takes the value of 100 if your leader chose Investment X, and x takes the value of 0 if your leader chose Investment Y.

To illustrate, suppose your leader has chosen Investment X. This means that $x = 100$ in the formula above, and your chance of receiving 10 ECU will be higher if your guess is higher. If you state 100 as your guess that the leader has chosen Investment X, then your chance of receiving 10 ECU will be $[1 - \left(\frac{100-100}{100}\right)^2] \times 100 = 100$. On the other hand, suppose your leader has chosen Investment Y instead, while your guess remains at 100. This means that $x = 0$ in the formula above, and your chance of receiving 10 ECU will be $[1 - \left(\frac{0-100}{100}\right)^2] \times 100 = 0$.

Here is another example:

Example 4. Suppose you guess 70 as the chance that your leader has chosen Investment X for the group. At the end of the experiment, the computer reveals that your leader has chosen Investment X for the group. Hence, your chance of receiving 10 ECU will be $[1 - \left(\frac{100-70}{100}\right)^2] \times 100 = 91$.

To determine whether you receive 10 ECU, the computer will randomly draw a number between 0 and 100 (including decimal points). If the number drawn by the computer is less than or equal to your chance of receiving 10 ECU as determined by the formula above, then you will receive 10 ECU. Otherwise, you will receive 0 ECU. Hence, in Example 4 above, if the number randomly drawn by the computer is less than or equal to 91, then you will receive 10 ECU. Otherwise, you will receive 0 ECU.

Payment for Experiment 2:

At the end of the experiment, if you are paid for Experiment 2, then the computer will randomly select one of the six rounds for payment. For the randomly chosen round:

1. If you are assigned the role of the leader, then you will be paid according to your investment decision in Stage 2 only.
2. If you are not assigned the role of the leader, then you will be paid according to your leader's investment decision in Stage 2, plus your decisions in Stage 3. The computer will randomly select one of the two questions in Stage 3, and you will be paid for your response to this question.

Summary

1. You will participate in six identical rounds in Experiment 2. At the beginning of each round, the computer will randomly match you to a new group with two other people. Each round consists of three stages.

2. In Stage 1, one group member will be assigned to be the leader of the group. There are four possible methods to determine who is assigned the role of the leader. You will be informed which method will be used to determine the leader at the beginning of each round.

   In Method 1, the computer will randomly assign one group member to be the leader.

   In Method 2, the group member who transferred the lowest amount to his/her matched partner in Experiment 1 will be assigned to be the leader.

   In Method 3, the group member who transferred the highest amount to his/her matched partner in Experiment 1 will be assigned to be the leader.

   In Method 4, you will be asked to indicate whether you prefer your leader to be someone who has transferred the highest or the lowest amount to his/her matched partner in Experiment 1. The computer will pick one of the decisions of the group members to implement. If your decision is implemented, then one of your other two group members will be appointed to be the leader based on your preference. Hence, you will not be appointed to be the leader if your decision is implemented.

   You will be asked to indicate your preferred leader for Method 4 once, at the beginning of Experiment 2. The computer will use the same decision whenever Method 4 is being used to determine the leader.

3. In Stage 2, you will be asked to make an investment decision, assuming that you have been assigned the role of the leader. The leader will be given an endowment of 300 ECU, and s/he will be asked to choose between two investment options that will affect the payoffs of all group members. Your decision will be implemented for your group only if you have been assigned the role of the leader for that round.

4. Investment X and Investment Y may be different in each round. In each round, the amount that you receive from each investment may be different, but you will always receive a higher payoff if the investment succeeds, and a lower payoff if it fails. The investment options will be shown on your computer screens.

5. In Stage 3, you will be asked to predict which investment your leader has chosen, assuming that you have not been assigned the role of the leader. You will be asked two questions.

   In Question 1, you will be asked to predict how likely it is in your opinion that the leader has chosen Investment X. You will need to choose a number between 0 and 100. A higher number means that you think the leader is more likely to have chosen Investment X.

   In Question 2, you are given additional information. Specifically, you will be asked the same question under two different scenarios: (i) suppose you are told that the investment has succeeded; and (ii) suppose you are told that the investment has failed. You should consider whether your guess of the leader's decision will be higher than, lower than, or the same as the one you stated in Question 1, given that you know the outcome of the investment chosen by your leader.

6. The payoff structure used to determine your payment in Stage 3 is designed such that it is in your best interest to report your true beliefs about your leader's decision.

7. At the end of the experiment, the computer will randomly select one of the six rounds for payment. For the randomly chosen round, if you are assigned the role of the leader, then you will be paid according to your decision in Stage 2. If you are not assigned the role of the leader, then you will be paid according to your leader's decision in Stage 2, as well as your decisions in Stage 3. The computer will randomly select one of the two questions in Stage 3 for payment.

If you have any questions, please raise your hand and an experimenter will come to you to answer your questions privately. Otherwise, please wait patiently for the experimenter to launch the practice questions on your computer screens. The purpose of these practice questions is to make sure that you understand the experiment. If you have any questions at any time, please raise your hand and an experimenter will come over to answer your questions privately.

Once everyone has completed the practice questions, we will proceed with one practice round for Experiment 2. The purpose of the practice round is to allow you to familiarize yourself with the decision screens. Your decisions in the practice round will not affect your payments for today's experiment. We will proceed with Experiment 2 once everyone has completed the practice round.

Practice Questions (Experiment 2)

1. I will be paid for the decisions in both experiments today. True/False [Ans: False]

2. We will participate in six identical rounds in Experiment 2. If we are paid for Experiment 2, then we will be paid for our decisions in one of the six rounds. True/False. [Ans: True]

3. We will participate in each round of Experiment 2 in groups of three. One group member will be assigned the role of the leader. True/False [Ans: True]

4. In Experiment 1, Player 1 chose to transfer 160 ECU to his/her matched partner, Player 2 chose to transfer 115 ECU to his/her matched partner, and Player 3 chose to transfer 160 ECU to his/her matched partner.

   Suppose the leader is appointed using Method 2. Which of the following is correct? [Ans: (b)]

   (a) Player 1 will be assigned the role of the leader.
   (b) Player 2 will be assigned the role of the leader.
   (c) Player 3 will be assigned the role of the leader.
   (d) Both Player 1 and Player 3 have an equal chance of being assigned the role of the leader.

5. In the above example, suppose the leader is appointed using Method 3. Which of the following is correct? [Ans: (d)]

   (a) Player 1 will be assigned the role of the leader.
   (b) Player 2 will be assigned the role of the leader.
   (c) Player 3 will be assigned the role of the leader.
   (d) Both Player 1 and Player 3 have an equal chance of being assigned the role of the leader.

6. Suppose the leader is appointed using Method 4. Suppose also that your preference for leadership appointment is randomly chosen by the computer to be implemented. Which of the following is correct? [Ans: (b)]

   (a) Depending on what I indicate as my preference of the appointed leader, I have a chance of being assigned the role of the leader.
   (b) Regardless of what I indicate as my preference of the appointed leader, I will definitely not be assigned the role of the leader.

7. Suppose the leader is appointed using Method 4. In Experiment 1, Player 1 chose to transfer 200 ECU to his/her matched partner, and Player 2 chose to transfer 85 ECU to his/her matched partner. Player 3 indicates that his/her preferred leader is someone who has transferred the highest amount to his/her matched partner in Experiment 1.

   Suppose Player 3's decision is randomly chosen by the computer to be implemented. Which of the following is correct? [Ans: (a)]

   (a) Player 1 will be assigned the role of the leader.
   (b) Player 2 will be assigned the role of the leader.
   (c) Both Player 1 and Player 2 have an equal chance of being assigned the role of the leader.

8. Which of the following is correct? [Ans: (b)]

   (a) The other group members will be informed of the investment chosen by the leader, but not the amount they have received from the investment.
   (b) The other group members will be informed of the amount they have received from the investment chosen by the leader, but not the investment chosen by him/her.
   (c) The other group members will be informed of the investment chosen by the leader, and the amount they have received from the investment.

9. Consider the investment options depicted in the figure below.

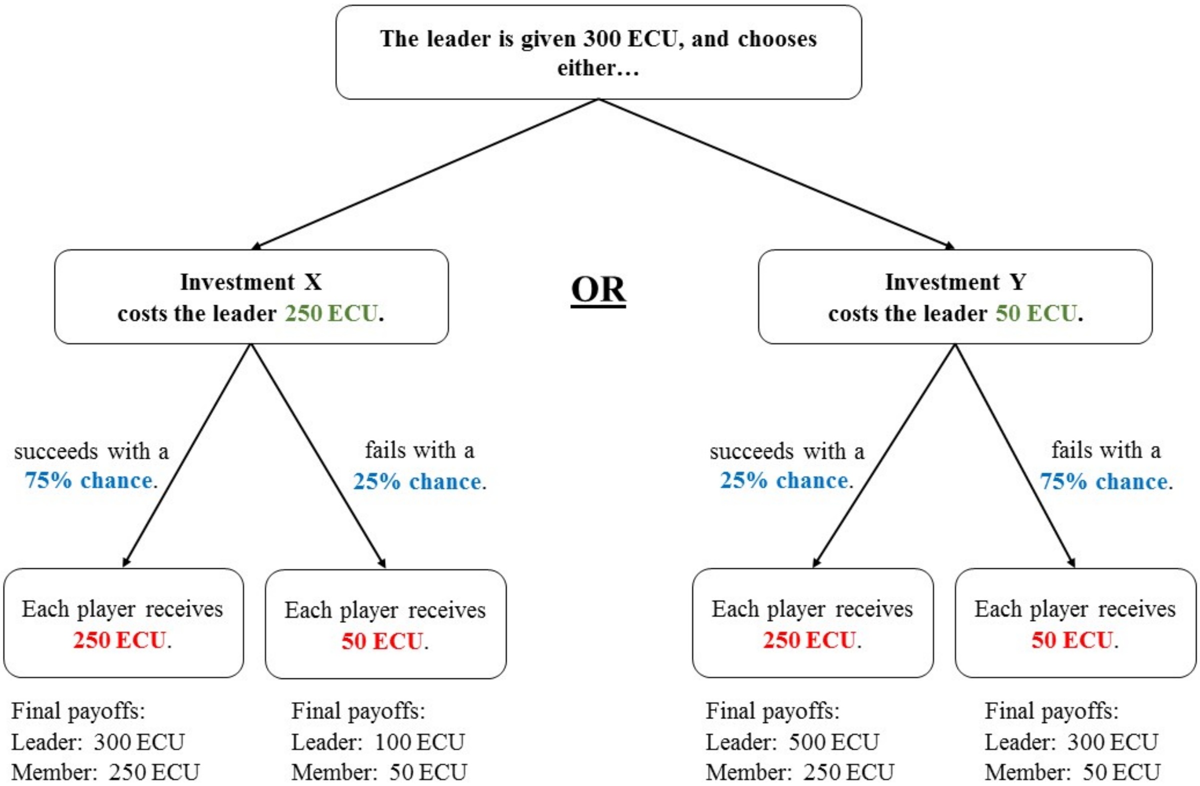


Figure A.2: Investment Options (Practice Question)

Suppose the leader chooses Investment X.

(a) At the end of the experiment, the computer randomly determines that the investment succeeds.

If you are not the leader, how many ECU will you receive from Stage 2 of Experiment 2? [Ans: 250 ECU]

(b) At the end of the experiment, the computer randomly determines that the investment fails.

If you are the leader, how many ECU will you receive from Stage 2 of Experiment 2? [Ans: 100 ECU]

10. Which of the following is true? [Ans: (c)]
    (a) I will be paid for my decision in Stage 3 of Experiment 2 regardless of whether I have been assigned the role of the leader or not.
    (b) I will be paid for my decision in Stage 3 of Experiment 2 only if I have been assigned the role of the leader.
    (c) I will be paid for my decision in Stage 3 of Experiment 2 only if I have not been assigned the role of the leader.

11. In Stage 3, I will be asked two questions. If I am paid for Stage 3 of Experiment 2, then I will be paid according to my answers to both questions. True/False [Ans: False]

12. Suppose you strongly believe that the leader of your group has chosen Investment Y. Which of the following statement is true? [Ans: (b)]
    (a) It is in my best interest to choose a higher number as my guess of "how likely is my leader to have chosen Investment X".
    (b) It is in my best interest to choose a lower number as my guess of "how likely is my leader to have chosen Investment X".
    (c) It is in my best interest to choose 50 as my guess of "how likely is my leader to have chosen Investment X".

A.2 Experiment 2 (Online)

Overview

Thank you for participating! You will receive $5 for completing today's experiment, and the instructions explain how you can make decisions and earn more money.

You will participate in two tasks today: Task 1 and Task 2. You will be paid for the decisions you make in either Task 1 or Task 2. Whether you will be paid for Task 1 or Task 2 will be randomly determined by the computer at the end of the session.

During the experiments, we will be using Experimental Currency Units (ECU). At the end of the session, we will convert the amount you earn into Australian Dollars (AUD) using the following conversion rate: 20 ECU = 1 AUD.

At the end of Task 2, you will be asked to fill out a brief questionnaire asking you some general questions. All of the decisions you make in today's session will remain anonymous, and you will never learn the identities of the other participants in today's session.

Task 1
(Treatment S)

You will participate in Task 1 in groups of three. The computer will randomly match you with two other people from today's experiment. You will stay in the same group for all of Task 1.

You will participate in three rounds of a decision task. Each round consists of two stages.

Stage 1: Investment decision by the group leader.

One group member will be randomly assigned by the computer to be the leader at the beginning of Task 1. In Stage 1, if you are the leader, you will be asked to make an investment decision for your group. Your role will remain the same for all three rounds of Task 1.

If you are the leader, you will be given an endowment of 300 ECU in each round. You will choose between two investment options that will affect both your payoff and the other group members' payoffs. Each investment can either fail or succeed. The two investment options have different chances of success/failure, as well as different costs to you.

Specifically, the two investment options are:

> Investment X: This investment costs you 250 ECU. It will succeed with a 75% chance and fail with a 25% chance.
>
> Investment Y: This investment costs you 50 ECU. It will succeed with a 25% chance and fail with a 75% chance.

The payoffs to the leader and each group member in this stage are calculated as follows:

1. Payoff to leader = 300 ECU − Cost of investment + Returns on investment
2. Payoff to each group member = Returns on investment

The table below shows an example where the return from each investment option is 200 ECU if it succeeds, and 0 ECU if it fails. These values are shown in **red**.

Investment Option	Cost to Leader	Investment		Payoff to each Member if investment:		Payoff to Leader if investment:	
		Succeeds	Fails	Succeeds	Fails	Succeeds	Fails
X	250	75%	25%	200	0	250 (= 300 − 250 + 200)	50 (= 300 − 250 + 0)
Y	50	25%	75%	200	0	450 (= 300 − 50 + 200)	250 (= 300 − 50 + 0)

Table A.1: Investment Options (Example of a Round)

Example. Suppose you are the leader, and you choose Investment X as shown in Table 1 above. Then, the investment costs you 250 ECU, and it will succeed with a 75% chance and fail with a 25% chance. If the investment succeeds, then you will receive (300 − 250 + 200) = 250 ECU and each of your group members will receive 200 ECU. Instead if the investment fails, then you will receive (300 − 250 + 0) = 50 ECU and each of your group members will receive 0 ECU.

Please note that in each round of Task 1, the returns from the investment options will be different. Within each round, both investments will provide the same high return if they succeed, and the same low return if they fail. However, the chance of failure and success will be different for different options (as stated above). Please pay attention to these values on the screen.

The other group members will never learn your investment decisions in the three rounds. At the end of the experiment, they will learn how much they will receive from the chosen investment, but they will not learn whether you chose Investment X or Y.

Stage 2: Beliefs of the other group members.

In Stage 2, if you are not assigned to be the leader of your group, you will be asked to predict how likely it is in your opinion that your leader has chosen Investment X.

The specific questions you will be asked are listed below.

Question 1

How likely do you think it is that your Leader has chosen Investment X? Specifically, what is the chance out of 100 that s/he has chosen Investment X?

In Question 2, you are given additional information. You are asked to evaluate the same question with this additional information.

Question 2(a)

Suppose you are informed that the investment chosen by your leader has succeeded. This gives you a high payoff.
Now consider whether your prediction will be higher than, lower than, or the same as the one you stated in Question 1. Specifically, given that the investment has succeeded, what is the chance out of 100 that s/he has chosen Investment X?

Question 2(b)

Suppose you are informed that the investment chosen by your leader has failed. This gives you a low payoff.
Now consider whether your prediction will be higher than, lower than, or the same as the one you stated in Question 1. Specifically, given that the investment has failed, what is the chance out of 100 that s/he has chosen Investment X?

For both questions, you will need to choose a number between 0 and 100. A higher number means that you think your leader is more likely to have chosen Investment X.

For your payment, the computer will randomly select one of these two questions and you will be paid for your response to this question. If Question 2 is chosen for payment, you will be paid for Question 2(a) if the investment has succeeded or Question 2(b) if it has failed.

To determine your payment, we use a procedure which has been used in many other studies. For the question randomly chosen by the computer, you receive either 200 ECU or 0 ECU. The closer your prediction is to the actual decision made by your leader, the higher is your chance of receiving 200 ECU. Hence, what is most important is that this procedure is designed such that it is in your best interest to report your true belief about the chance that your leader has chosen Investment X.

[The exact details of how your payment will be determined are available here (link to separate document) if you are interested, but it is not necessary for you to read these notes.]

Payment for Task 1:

At the end of the experiment, if you are paid for Task 1, then the computer will randomly select one of the three rounds for payment. For the randomly chosen round:

1. **If you are the leader**, then you will be paid according to your investment decision in Stage 1 only.
2. **If you are not the leader**, then you will be paid either according to your leader's investment decision in Stage 1, or your predictions in Stage 2, but not both. The computer will randomly determine which one you will be paid for.

Summary

1. In Task 1, the computer will randomly divide you into groups of three. One group member will be randomly assigned by the computer to be the leader of the group. You will stay in the same group and role for all of Task 1. You will be informed whether you are assigned to be the leader at the beginning of Task 1.

2. You will participate in three rounds in Task 1. Each round consists of two stages.

3. In Stage 1, if you are the leader, you will be asked to make an investment decision. You will be given an endowment of 300 ECU in each round, and you will choose between two investment options that will affect both your payoffs and the other group members' payoffs.

4. The returns from Investment X and Investment Y will be different in each round. However, within each round, both investments always provide the same high return if they succeed and the same low return if they fail.

5. In Stage 2, if you are not the leader, you will be asked to predict how likely it is in your opinion that your leader has chosen Investment X. The payoff structure used to determine your payment in Stage 2 is designed such that it is in your best interest to report your true belief about your leader's decision.

6. At the end of the experiment, the computer will randomly select one of the three rounds for payment. For the randomly chosen round:
   (a) If you are the leader, then you will be paid according to your decision in Stage 1.
   (b) If you are not the leader, then you will be paid either according to your leader's decision in Stage 1, or your predictions of your leader's decision in Stage 2.

You have arrived at the end of the instructions for Task 1.

Please return to the experiment and click the button on the screen to start the practice questions.

Task 1

(Treatment D)

You will participate in Task 1 in groups of three. The computer will randomly match you with two other people from today's experiment. You will stay in the same group for all of Task 1.

You will participate in three rounds of a decision task. Each round consists of two stages.

Stage 1: Investment decision by the group leader.

One group member will be randomly assigned by the computer to be the leader at the beginning of Task 1. You will be informed of your role at the end of the experiment. In Stage 1, you will be asked to make an investment decision assuming that you are the leader of your group. Your decision will be implemented at the end of the experiment if you are the leader of your group. Your role will remain the same for all three rounds of Task 1.

If you are the leader, you will be given an endowment of 300 ECU in each round. You will choose between two investment options that will affect both your payoff and the other group members' payoffs. Each investment can either fail or succeed. The two investment options have different chances of success/failure, as well as different costs to you.

Specifically, the two investment options are:

> Investment X: This investment costs you 250 ECU. It will succeed with a 75% chance and fail with a 25% chance.
>
> Investment Y: This investment costs you 50 ECU. It will succeed with a 25% chance and fail with a 75% chance.

The payoffs to the leader and each group member in this stage are calculated as follows:

1. Payoff to leader = 300 ECU − Cost of investment + Returns on investment
2. Payoff to each group member = Returns on investment

The table below shows an example where the return from each investment option is 200 ECU if it succeeds, and 0 ECU if it fails. These values are shown in **red**.

Investment Option	Cost to Leader	Investment		Payoff to each Member if investment:		Payoff to Leader if investment:	
		Succeeds	Fails	Succeeds	Fails	Succeeds	Fails
X	250	75%	25%	200	0	250 (= 300 − 250 + 200)	50 (= 300 − 250 + 0)
Y	50	25%	75%	200	0	450 (= 300 − 50 + 200)	250 (= 300 − 50 + 0)

Table A.2: Investment Options (Example of a Round)

Example. Suppose you are the leader, and you choose Investment X as shown in Table 1 above. Then, the investment costs you 250 ECU, and it will succeed with a 75% chance and fail with a 25% chance. If the investment succeeds, then you will receive (300 − 250 + 200) = 250 ECU and each of your group members will receive 200 ECU. Instead if the investment fails, then you will receive (300 − 250 + 0) = 50 ECU and each of your group members will receive 0 ECU.

Please note that in each round of Task 1, the returns from the investment options will be different. Within each round, both investments will provide the same high return if they succeed, and the same low return if they fail. However, the chance of failure and success will be different for different options (as stated above). Please pay attention to these values on the screen.

If, at the end of the experiment, your decisions are implemented as the leader of your group, the other group members will never learn your investment decisions in the three rounds. At the end of the experiment, they will learn how much they will receive from the chosen investment, but they will not learn whether you chose Investment X or Y.

Stage 2: Beliefs of the other group members.

In Stage 2, you will be asked to predict how likely it is in your opinion that the leader has chosen Investment X, <u>assuming that someone else in your group has been assigned the role of the leader</u>.

The specific questions you will be asked are listed below.

<u>Question 1</u>

How likely do you think it is that your Leader has chosen Investment X? Specifically, what is the chance out of 100 that s/he has chosen Investment X?

In Question 2, you are given additional information. You are asked to evaluate the same question with this additional information.

<u>Question 2(a)</u>

Suppose you are informed that the investment chosen by your leader has succeeded. This gives you a high payoff.
Now consider whether your prediction will be higher than, lower than, or the same as the one you stated in Question 1. Specifically, given that the investment has succeeded, what is the chance out of 100 that s/he has chosen Investment X?

<u>Question 2(b)</u>

Suppose you are informed that the investment chosen by your leader has failed. This gives you a low payoff.
Now consider whether your prediction will be higher than, lower than, or the same as the one you stated in Question 1. Specifically, given that the investment has failed, what is the chance out of 100 that s/he has chosen Investment X?

For both questions, you will need to choose a number between 0 and 100. <u>A higher number means that you think your leader is more likely to have chosen Investment X.</u>

For your payment, the computer will randomly select one of these two questions and you will be paid for your response to this question. If Question 2 is chosen for payment, you will be paid for Question 2(a) if the investment has succeeded or Question 2(b) if it has failed.

To determine your payment, we use a procedure which has been used in many other studies. For the question randomly chosen by the computer, you receive either 200 ECU or 0 ECU. The closer your prediction is to the actual decision made by your leader, the higher is your chance of receiving 200 ECU. Hence, what is most important is that this procedure is designed such that it is in your best interest to report your true belief about the chance that your leader has chosen Investment X.

[The exact details of how your payment will be determined are available here (link to separate document) if you are interested, but it is not necessary for you to read these notes.]

Payment for Task 1:

At the end of the experiment, if you are paid for Task 1, then the computer will randomly select one of the three rounds for payment. For the randomly chosen round:

1. **If you are the leader**, then you will be paid according to your investment decision in Stage 1 only.

2. **If you are not the leader**, then you will be paid either according to your leader's investment decision in Stage 1, or your predictions in Stage 2, but not both. The computer will randomly determine which one you will be paid for.

Summary

1. In Task 1, the computer will randomly divide you into groups of three. One group member will be randomly assigned by the computer to be the leader of the group. You will stay in the same group and role for all of Task 1. You will be informed whether you are assigned to be the leader at the end of the experiment.

2. You will participate in three rounds in Task 1. Each round consists of two stages.

3. In Stage 1, you will be asked to make an investment decision, assuming that you are the leader. You will be given an endowment of 300 ECU in each round, and you will choose between two investment options that will affect both your payoffs and the other group members' payoffs. At the end of the experiment, your decision will be implemented for your group only if you are the leader of your group.

4. The returns from Investment X and Investment Y will be different in each round. However, within each round, both investments always provide the same high return if they succeed and the same low return if they fail.

5. In Stage 2, you will be asked to predict how likely it is in your opinion that your leader has chosen Investment X, assuming that someone else in your group has been assigned the role of the leader. The payoff structure used to determine your payment in Stage 2 is designed such that it is in your best interest to report your true belief about your leader's decision.

6. At the end of the experiment, the computer will randomly select one of the three rounds for payment. For the randomly chosen round:
   (a) If you are the leader, then you will be paid according to your decision in Stage 1.
   (b) If you are not the leader, then you will be paid either according to your leader's decision in Stage 1, or your predictions of your leader's decision in Stage 2.

You have arrived at the end of the instructions for Task 1.

Please return to the experiment and click the button on the screen to start the practice questions.

Task 2

You will participate in Task 2 in groups of two. The computer will randomly match you with one other person in today's session.

Each of you is given an endowment of 300 ECU, and you are asked to divide this amount between yourself and the person you are matched with.

At the end of today's session, if Task 2 is picked for payment, then you will be paid either according to your decision or according to the decision made by your randomly matched partner. The computer will randomly determine whose allocation decision will be implemented.

Example. Suppose you choose to divide your endowment by keeping 200 ECU for yourself and giving 100 ECU to your matched partner. Your matched partner decides to keep 130 ECU and give 170 ECU to you. If, at the end of the experiment, the computer randomly determines that it is the allocation of your matched partner that gets implemented, then your payment will be 170 ECU and your matched partner's payment will be 130 ECU.

Are there any questions? If not, we will proceed with Task 2.

Practice Questions (Task 1)

(Treatments S and D)

Note: Each question is shown on a separate screen on oTree.

1. I will be paid for the decisions in both tasks today. True/False [Ans: False]
2. Task 1 has three rounds. If Task 1 is chosen for payment, then I will be paid for the decisions in one of the three rounds. True/False [Ans: True]
3. (Treatment S) We will participate in Task 1 in groups of three. One group member will be assigned the role of the leader. I will remain in the same group and role for all rounds of Task 1. True/False [Ans: True]

   (Treatment D) We will participate in Task 1 in groups of three. One group member will be assigned the role of the leader. I will remain in the same group for all rounds of Task 1. True/False [Ans: True]
4. Which of the following is correct? [Ans: (b)]
   (a) The other group members will be informed of the investment chosen by the leader, but not the amount they have received from the investment.
   (b) The other group members will be informed of the amount they have received from the investment chosen by the leader, but not the investment chosen by him/her.
   (c) The other group members will be informed of the investment chosen by the leader, and the amount they have received from the investment.
5. Consider the investment options given in the table below.

Investment Option	Cost to Leader	Investment		Payoff to each Member if investment:		Payoff to Leader if investment:	
		Succeeds	Fails	Succeeds	Fails	Succeeds	Fails
X	250	75%	25%	250	50	300 (= 300 − 250 + 250)	100 (= 300 − 250 + 50)
Y	50	25%	75%	250	50	500 (= 300 − 50 + 250)	300 (= 300 − 50 + 50)

Suppose the leader chooses **Investment X**.

At the end of the experiment, the computer randomly determines that the investment **succeeds**.

If you are **not the leader**, how many ECU will you receive from Stage 1? [Ans: 250]

6. Consider the investment options given in the table below.

Investment Option	Cost to Leader	Investment		Payoff to each Member if investment:		Payoff to Leader if investment:	
		Succeeds	Fails	Succeeds	Fails	Succeeds	Fails
X	250	75%	25%	250	50	300 (= 300 − 250 + 250)	100 (= 300 − 250 + 50)
Y	50	25%	75%	250	50	500 (= 300 − 50 + 250)	300 (= 300 − 50 + 50)

Suppose the leader chooses **Investment X**.

At the end of the experiment, the computer randomly determines that the investment **fails**.

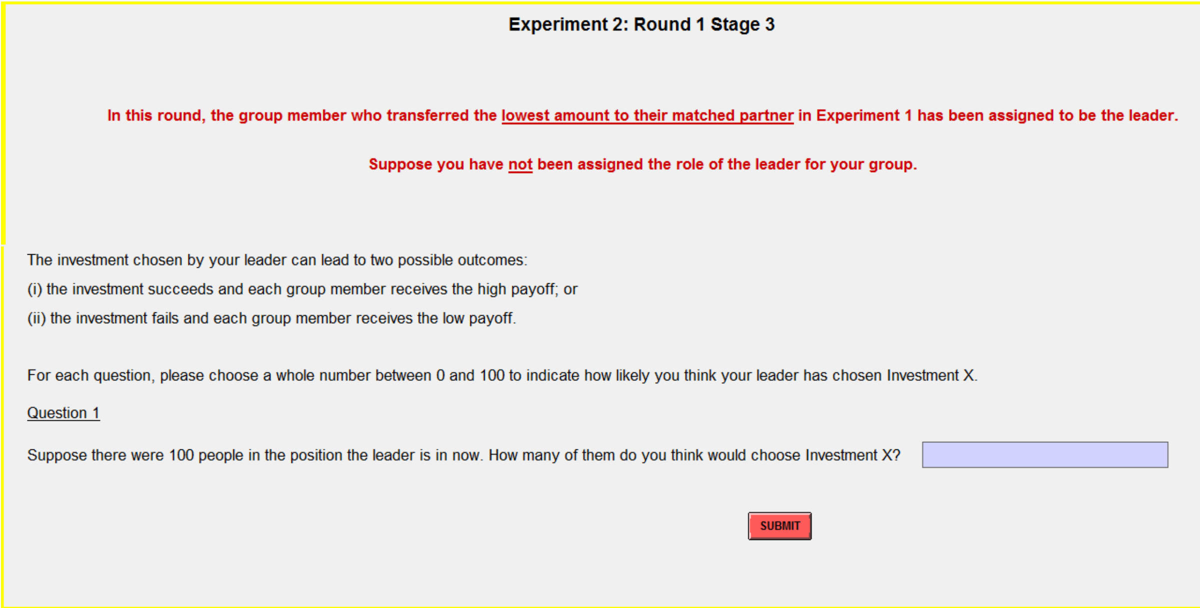
If you are **the leader**, how many ECU will you receive from Stage 1? [Ans: 100]

7. Which of the following is true? [Ans: (c)]
   (a) If I am the leader, then I will be paid for my decisions in Stage 2.
   (b) If I am not the leader, then I will be paid for BOTH my leader's decision in Stage 1 AND my decisions in Stage 2.
   (c) If I am not the leader, then I will be paid for EITHER my leader's decision in Stage 1 OR my decisions in Stage 2.

8. If I am not the leader and I am paid for Stage 2, then I will be paid according to my answers to both Question 1 and Question 2. True/False [Ans: False]

9. (Treatment S) Suppose you strongly believe that the leader of your group has chosen Investment Y.
   (Treatment D) Suppose you are not the leader. You strongly believe that the leader of your group has chosen Investment Y.
   Which of the following statement is true? [Ans: (b)]
   (a) It is in my best interest to choose a high number as my prediction of the chance that my leader has chosen investment X.
   (b) It is in my best interest to choose a low number as my prediction of the chance that my leader has chosen investment X.
   (c) It is in my best interest to choose 50 as my prediction of the chance that my leader has chosen investment X.

B Screenshots for belief elicitation task in Experiment 1



Experiment 2: Round 1 Stage 3

In this round, the group member who transferred the lowest amount to their matched partner in Experiment 1 has been assigned to be the leader.

Suppose you have not been assigned the role of the leader for your group.

The investment chosen by your leader can lead to two possible outcomes:
(i) the investment succeeds and each group member receives the high payoff; or
(ii) the investment fails and each group member receives the low payoff.

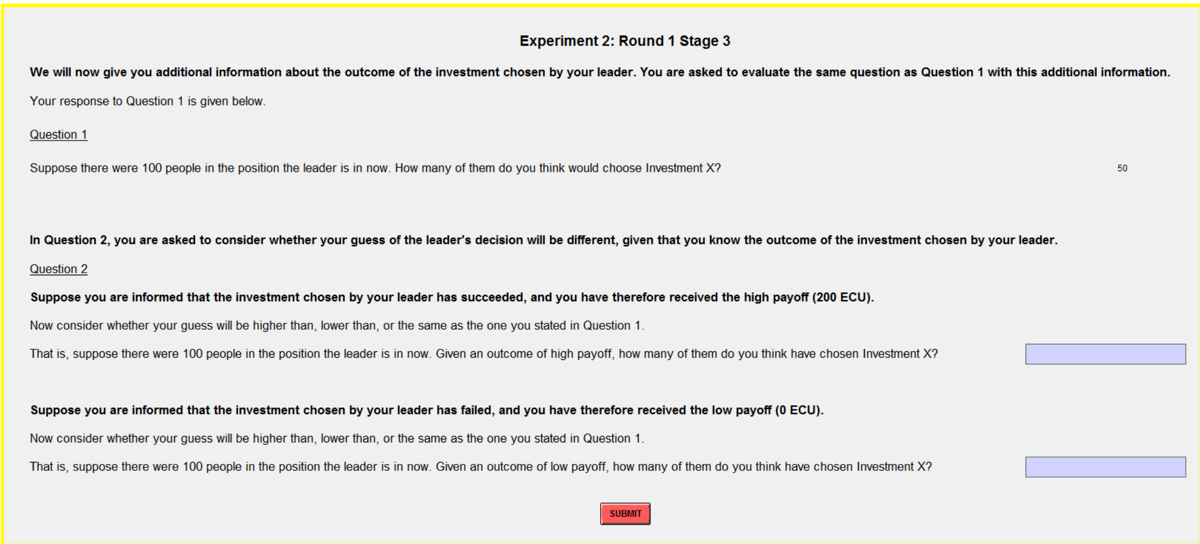
For each question, please choose a whole number between 0 and 100 to indicate how likely you think your leader has chosen Investment X.

Question 1

Suppose there were 100 people in the position the leader is in now. How many of them do you think would choose Investment X?

SUBMIT

(a) Interim Belief



Experiment 2: Round 1 Stage 3

We will now give you additional information about the outcome of the investment chosen by your leader. You are asked to evaluate the same question as Question 1 with this additional information.

Your response to Question 1 is given below.

Question 1

Suppose there were 100 people in the position the leader is in now. How many of them do you think would choose Investment X? 50

In Question 2, you are asked to consider whether your guess of the leader's decision will be different, given that you know the outcome of the investment chosen by your leader.

Question 2

Suppose you are informed that the investment chosen by your leader has succeeded, and you have therefore received the high payoff (200 ECU).

Now consider whether your guess will be higher than, lower than, or the same as the one you stated in Question 1.

That is, suppose there were 100 people in the position the leader is in now. Given an outcome of high payoff, how many of them do you think have chosen Investment X?

Suppose you are informed that the investment chosen by your leader has failed, and you have therefore received the low payoff (0 ECU).

Now consider whether your guess will be higher than, lower than, or the same as the one you stated in Question 1.

That is, suppose there were 100 people in the position the leader is in now. Given an outcome of low payoff, how many of them do you think have chosen Investment X?

SUBMIT

(b) Posterior Beliefs

Figure B.1: Decision screens – Elicitation of beliefs

C Derivation of Hypothesis 1

A member of the group is appointed to be the DM under one of four possible appointment mechanisms, $\Psi \in \{RA, LA, HA, GA\}$. We are interested in how members form interim beliefs about their DM's type under each appointment mechanism, given that they have been informed that someone else in the group is the DM.

Random appointment (RA). Each member has an equal chance of being appointed as the DM. This implies that

$$\mu_i^{RA} = \Pr(\beta \geq \beta^*) = 1 - F(\beta^*). \tag{C.1}$$

Appointment of lowest type (LA). The member with the lowest β is appointed to be the DM. Consider member i of type β_i who is informed that someone else in the group has been appointed to be the DM under this mechanism. Hence, s/he knows that the DM has type $\beta \leq \beta_i$ as otherwise s/he would have been appointed to be the DM. Denote the minimum of the remaining $N-1$ members' types as $\beta^{\min}$.

Given this, there are two possible cases. First, if $\beta_i < \beta^*$, then it must be that $\mu_i^{LA} = 0$ since the DM has type $\beta \leq \beta_i < \beta^*$. Second, if $\beta_i \geq \beta^*$, then the probability that the appointed DM is of type $\beta \geq \beta^*$ is given by

$$\begin{aligned}
\Pr(\beta^{\min} \geq \beta^* | \beta^{\min} < \beta_i) &= 1 - \Pr(\beta^{\min} < \beta^* | \beta^{\min} < \beta_i) \\
&= 1 - \frac{\Pr(\beta^{\min} < \beta^*)}{\Pr(\beta^{\min} < \beta_i)} \quad (\text{since } \beta^* \leq \beta_i) \\
&= 1 - \frac{1 - [1 - F(\beta^*)]^{N-1}}{1 - [1 - F(\beta_i)]^{N-1}}.
\end{aligned}$$

Hence, for member i,

$$\mu_i^{LA} = \begin{cases} 0 & \text{if } \beta_i < \beta^*, \\ \frac{[1-F(\beta^*)]^{N-1} - [1-F(\beta_i)]^{N-1}}{1-[1-F(\beta_i)]^{N-1}} & \text{if } \beta_i \geq \beta^*. \end{cases} \tag{C.2}$$

Clearly $\mu_i^{LA} \leq \mu_i^{RA}$ for $\beta_i < \beta^*$, which holds as an equality if $\beta^* = 1$. For $\beta_i \geq \beta^*$, $\mu_i^{LA} - \mu_i^{RA} = \frac{[1-F(\beta^*)]^{N-1} - [1-F(\beta_i)]^{N-1} - [1-F(\beta^*)]\{1-[1-F(\beta_i)]^{N-1}\}}{1-[1-F(\beta_i)]^{N-1}}$. The denominator is ≥ 0. The numerator can be simplified to give $[1-F(\beta^*)]^{N-1} - [1-F(\beta^*)] - F(\beta^*)[1-F(\beta_i)]^{N-1}$, which is ≤ 0 since $[1-F(\beta^*)]^{N-1} \leq [1-F(\beta^*)]$. Hence, $\mu_i^{LA} \leq \mu_i^{RA}$ for $\beta_i \geq \beta^*$.

Appointment of highest type (HA). The individual with the highest β is appointed to be the DM. Consider member i of type β_i who is informed that someone else in the group has been appointed to be the DM under this mechanism. Hence, s/he knows that the DM has type $\beta \geq \beta_i$ as otherwise s/he would have been appointed as the DM. Denote the maximum of the remaining $N-1$ members' types as $\beta^{\max}$.

Given this, there are two possible cases. First, if $\beta_i \geq \beta^*$, then it must be that $\mu_i^{HA} = 1$ since the DM is of type $\beta \geq \beta_i \geq \beta^*$. Second, if $\beta_i < \beta^*$, then the probability that the appointed DM is of type $\beta \geq \beta^*$ is given by

$$\begin{aligned}
\Pr(\beta^{\max} \geq \beta^* | \beta^{\max} \geq \beta_i) &= \frac{\Pr(\beta^{\max} \geq \beta^*)}{\Pr(\beta^{\max} \geq \beta_i)} \quad (\text{since } \beta^* > \beta_i) \\
&= \frac{1 - F(\beta^*)^{N-1}}{1 - F(\beta_i)^{N-1}}.
\end{aligned}$$

Hence, for member i,

$$\mu_i^{HA} = \begin{cases} 1 & \text{if } \beta_i \geq \beta^*, \\ \frac{1-F(\beta^*)^{N-1}}{1-F(\beta_i)^{N-1}} & \text{if } \beta_i < \beta^*. \end{cases} \tag{C.3}$$

Clearly $\mu_i^{HA} \geq \mu_i^{RA}$ for $\beta_i \geq \beta^*$, which holds as an equality if $\beta^* = 0$. For $\beta_i < \beta^*$, $\mu_i^{HA} - \mu_i^{RA} \geq 0$ since $1 - F(\beta^*)^{N-1} \geq 1 - F(\beta^*)$ and $1 - F(\beta_i)^{N-1} \leq 1$. Hence, $\mu_i^{HA} \geq \mu_i^{RA}$ for $\beta_i < \beta^*$ also.

Group appointment (GA). All members indicate how they would like their DM to be appointed. Specifically, they may choose to appoint as DM: (i) the lowest-type member; (ii) the highest-type member; or (iii) a randomly picked member. One of the group members' appointment decisions is randomly chosen to be implemented and the DM is appointed from the remaining group members based on this individual's preference.

It is trivial to see that all members will prefer to have the highest type appointed as the DM regardless of their own type. Intuitively, this is because it increases the chance that the appointed DM is of type $\beta \geq \beta^*$ and chooses a high effort level, leading to higher expected payoffs for the members.

Consider member i of type β_i who is informed that someone else in the group has been appointed to be the DM. There are two possible cases. First, if member i's appointment decision is implemented, then the probability that the DM is of type $\beta \geq \beta^*$ depends on the probability that at least one of the other $N-1$ group members is of type $\geq \beta^*$. This is given by $1 - F(\beta^*)^{N-1}$. Second, if member i's appointment decision is not implemented, then s/he knows that the DM is of type $\beta \geq \beta_i$ as otherwise s/he would have been appointed to be the DM. Specifically, the DM's type is given by the maximum of the remaining $N-2$ members' types (excluding member i and the member whose decision is implemented). The derivation of the probability that the DM is of type $\beta \geq \beta^*$ under this scenario is similar to that of mechanism HA with $N-2$ other members.

We next evaluate member i's posterior belief that his/her appointment decision has been implemented, given the information that someone else in the group has been appointed to be the DM. Using Bayes' rule, we get $\frac{\frac{1}{N}\cdot(1)}{\frac{1}{N}\cdot(1)+\frac{N-1}{N}[1-F(\beta_i)^{N-2}]} = \frac{1}{1+(N-1)[1-F(\beta_i)^{N-2}]}$. The numerator is the product of the prior probability that member i's appointment decision is implemented ($\frac{1}{N}$) and the probability that s/he is not assigned to be the DM conditional on having his/her decision implemented. Conditional on member i's decision being implemented, s/he does not become the DM with certainty. The denominator is the probability that member i is not appointed to be the DM. The first term is the same as the numerator. The second term is the product of the prior probability that member i's appointment decision is not implemented ($\frac{N-1}{N}$) and the probability that s/he is not assigned to be the DM conditional on not having his/her appointment decision implemented. Conditional on member i's appointment decision not being implemented, the probability that member i is not appointed to be the DM is $1 - F(\beta_i)^{N-2}$. This is the probability that at least someone else in the group (other than both member i and the member whose decision is implemented) has type $\beta \geq \beta_i$ and is therefore appointed to be the DM.

Putting all these together, we have for member i,

$$\mu_i^{GA} = A \times \left[1 - F(\beta^*)^{N-1}\right] + (1-A) \times \begin{cases} 1 & \text{if } \beta_i \geq \beta^*, \\ \frac{1-F(\beta^*)^{N-2}}{1-F(\beta_i)^{N-2}} & \text{if } \beta_i < \beta^*, \end{cases} \tag{C.4}$$

where $A \equiv \frac{1}{1+(N-1)[1-F(\beta_i)^{N-2}]}$.

We would like to show that $\mu_i^{GA} \geq \mu_i^{RA}$. Note that μ_i^{GA} is a convex combination of two terms since $A \leq 1$. For both $\beta_i \geq \beta^*$ and $\beta_i < \beta^*$, these two terms are $\geq \mu_i^{RA}$ for $\beta^* > 0$. For $\beta^* = 0$, $\mu_i^{GA} = \mu_i^{RA}$.

Next, we would like to show that $\mu_i^{GA} \leq \mu_i^{HA}$. Again, since μ_i^{GA} is a convex combination of two terms, it is sufficient to show that these two terms are $\leq \mu_i^{HA}$. This is clearly the case for $\beta_i \geq \beta^*$. For $\beta_i < \beta^*$, we need to show that $\frac{1-F(\beta^*)^{N-2}}{1-F(\beta_i)^{N-2}} \leq \frac{1-F(\beta^*)^{N-1}}{1-F(\beta_i)^{N-1}}$. This is equivalent to showing that

$$\frac{[1-F(\beta^*)^{N-1}][1-F(\beta_i)^{N-2}] - [1-F(\beta^*)^{N-2}][1-F(\beta_i)^{N-1}]}{[1-F(\beta_i)^{N-2}][1-F(\beta_i)^{N-1}]} \geq 0.$$

The denominator is ≥ 0. Let $x \equiv F(\beta^*)$ and $y \equiv F(\beta_i)$ with $x > y$ since $\beta^* > \beta_i$. Then, the numerator becomes $(1-x^{N-1})(1-y^{N-2}) - (1-x^{N-2})(1-y^{N-1})$. Simplifying gives us

$$x^{N-2} - x^{N-1} + y^{N-1} - y^{N-2} + x^{N-1}y^{N-2} - x^{N-2}y^{N-1} \tag{C.5}$$

Hence, for the numerator to be ≥ 0, we need to show the following:

Claim: $x^{N-2} - x^{N-1} + y^{N-1} - y^{N-2} + x^{N-1}y^{N-2} - x^{N-2}y^{N-1} \geq 0$ for $x, y \in [0,1]$, $x > y$, and $N > 2$.

Proof. The proof is by induction. Let $x \equiv \alpha y$, $\alpha > 1$. Then, (C.5) becomes

$$(\alpha y)^{N-2} - (\alpha y)^{N-1} + y^{N-1} - y^{N-2} + \alpha^{N-1} y^{2N-3} - \alpha^{N-2} y^{2N-3} \tag{C.6}$$

Consider first $N = 3$. (C.6) becomes

$$\alpha y - (\alpha y)^2 + y^2 - y + \alpha^2 y^3 - \alpha y^3 = y(\alpha - 1)(1 - y)(1 - \alpha y)$$

which is ≥ 0 since $y \in [0, 1]$, $\alpha y = x \in [0, 1]$, and $\alpha > 1$. Now suppose (C.6) ≥ 0 for some $N = k$. Rearranging (C.6), we have

$$\alpha^{k-2} y^{k-2}(1 - \alpha y) + \alpha^{k-2} y^{2k-3}(\alpha - 1) \geq y^{k-2} - y^{k-1}. \tag{C.7}$$

Next consider $N = k+1$. (C.6) becomes $\alpha^{k-1} y^{k-1}(1-\alpha y) + y^k - y^{k-1} + \alpha^{k-1} y^{2k-1}(\alpha - 1)$, which is equal to

$$y \left[\alpha^{k-1} y^{k-2}(1 - \alpha y) + \alpha^{k-1} y^{2k-2}(\alpha - 1) + y^{k-1} - y^{k-2}\right]. \tag{C.8}$$

We want to show that this expression is ≥ 0 given that (C.7) holds. Since $y \geq 0$, this is equivalent to showing the terms inside the brackets are ≥ 0, or

$$\alpha^{k-1} y^{k-2}(1 - \alpha y) + \alpha^{k-1} y^{2k-2}(\alpha - 1) \geq y^{k-2} - y^{k-1}. \tag{C.9}$$

Note that the RHS of (C.9) is the same as the RHS of (C.7) and is ≥ 0. To conclude the proof, we show that the LHS of (C.9) is $\geq$ the LHS of (C.7). Note that this is equivalent to showing

$$(\alpha - 1)(1 - \alpha y)\alpha^{k-2}(y^{k-2} - y^{(k-2)+(k-1)}) \geq 0,$$

which holds because $\alpha > 1$, $\alpha y \in [0, 1]$, and $y^{k-2} \geq y^{(k-2)+(k-1)}$. Hence, (C.8) is ≥ 0 if (C.7) holds.

■

D Additional analysis

D.1 Analysis with Game 1 treatments only

This section presents the analyses with the Game 1 treatments only. We show that Results 1 and 2 hold with the exclusion of the Game 0 treatments.

Table D.1 presents marginal-effects estimates from a probit model for the relationship between the subjects' decisions as DMs in the investment task and their dictator game behavior. The estimates in the table reveal that a DM who transfers 1% more of their endowment to their matched partner in the dictator game is 0.4% more likely to choose e_H in the investment task on average, and this effect is statistically significant (p-value < 0.001). Hence, we conclude that the dictator game is a good proxy for an individual's type β_i even when we consider only the Game 1 treatments.

Table D.1: Regression of DM's effort choice (Game 1)

Variables	Dependent variable: =1 if DM chooses e_H (1)
% endowment transferred in DG	0.004***
	(0.001)
% endowment invested in RT	−0.001
	(0.001)
Treatment LA	−0.045
	(0.028)
Treatment HA	0.048
	(0.030)
Treatment GA	0.039
	(0.028)
Order Effects	Y
Observations	1,088
# subjects (clusters)	272

Marginal effects of probit model reported. Robust standard errors in parentheses. Standard errors are clustered at the subject level.
DG: Dictator Game; RT: Risk Task.
*** p<0.01, ** p<0.05, * p<0.10.

Table D.2 presents OLS estimates for the regressions of interim beliefs against the treatment variables. Similar to the main analysis in the paper, we control for order effects in columns (1) and (3) and individual fixed effects in columns (2) and (4). Treatment RA is the comparison group in all the specifications. Overall, the coefficient estimates reveal that Result 1 is robust to the exclusion of the Game 0 treatments. In particular, group members respond to the appointment mechanism in their interim beliefs in the Game 1 treatments.

Table D.2: Regression of members' interim belief (Game 1)

	Dependent variable: Interim belief			
Variables	(1)	(2)	(3)	(4)
Treatment LA	-13.074^{***}	-13.074^{***}	-11.989^{***}	-12.375^{***}
	(1.484)	(1.482)	(1.465)	(1.425)
Treatment HA	9.787^{***}	9.787^{***}	8.702^{***}	9.088^{***}
	(1.397)	(1.396)	(1.352)	(1.332)
Treatment GA	2.717^{**}	2.717^{**}	1.813	2.135^{*}
	(1.355)	(1.354)	(1.277)	(1.265)
Chooses high effort as DM			24.584^{***}	15.832^{***}
			(1.960)	(1.984)
% endowment invested in RT	-0.086^{*}		-0.055	
	(0.045)		(0.039)	
Constant	55.525^{***}	45.938^{***}	45.400^{***}	41.572^{***}
	(3.990)	(0.812)	(3.636)	(0.890)
Order Effects	Y	N	Y	N
Individual FE	N	Y	N	Y
Observations	1,088	1,088	1,088	1,088
# subjects (clusters)	272	272	272	272
R-squared	0.111	0.233	0.278	0.306
Test of HA = GA				
test statistic	5.202	5.209	5.341	5.461
p-value	0.000	0.000	0.000	0.000

Robust standard errors clustered at the subject level in parentheses. For all regressions, treatment RA is the reference treatment.
RT: Risk Task.
*** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table D.3 presents the results from the OLS estimation of equation (2). Similar to the main analysis in the paper, we drop the inconsistent and non-updaters in the analysis. We find that Result 2 is also robust to the exclusion of the Game 0 treatments. Within the Game 1 treatments, members suffer from base-rate neglect relative to a Bayesian (test of $\delta = 1$: p-value < 0.001), attribute good outcomes more to luck than a Bayesian would (test of $\gamma_G = 1$: p-value < 0.001), and treat bad outcomes like a Bayesian (test of $\gamma_B = 1$: p-value $= 0.492$). Consequently, they tend to attribute good (bad) outcomes more to the DM's luck (decision), i.e., $\gamma_G < \gamma_B$ (p-value $= 0.012$). While members exhibit similar biases in their updating behavior across all the appointment mechanisms, unlike Table 4, the asymmetry in the attribution of outcomes is now marginally statistically insignificant in treatment LA (p-value $= 0.103$) and statistically significant in treatment HA (p-value $= 0.028$).

Table D.3: Regression of members' posterior beliefs (Game 1)

	Dependent variable: Logit(posterior)				
	(1)	(2)	(3)	(4)	(5)
Variables	Pooled	RA	LA	HA	GA
δ : logit(interim belief)	0.733***	0.764***	0.793***	0.771**	0.529***
	(0.049)	(0.071)	(0.057)	(0.093)	(0.135)
γ_G : Good outcome $\times$ logit(p)	0.742***	0.744***	0.728***	0.752***	0.798**
	(0.060)	(0.089)	(0.078)	(0.094)	(0.098)
γ_B : Bad outcome $\times$ logit($1-p$)	0.948	0.932	0.937	0.994	0.876
	(0.076)	(0.092)	(0.119)	(0.090)	(0.114)
Observations	1,640	410	410	410	410
# subjects (clusters)	205	205	205	205	205
R-squared	0.636	0.686	0.741	0.613	0.421
Test of $\gamma_G = \gamma_B$					
test statistic	−2.522	−1.588	−1.637	−2.218	−0.512
p-value	0.012	0.114	0.103	0.028	0.609

Robust standard errors clustered at the subject level in parentheses.
This analysis includes only the Game 1 treatments but includes subjects classified as inconsistent or non-updaters.
*** p<0.01, ** p<0.05, * p<0.10. Null hypothesis is coefficient = 1.

D.2 IV regression of posterior beliefs

Table D.4 presents the results from the IV estimation of equation (2). We use the appointment mechanisms as instruments for the logit of members' interim beliefs.[2] The conclusions from the IV estimates are similar to those obtained from the OLS estimates in column (1) of Table 4. Specifically, we find that members suffer from base-rate neglect relative to a Bayesian (test of $\delta = 1$: p-value < 0.001). Moreover, members attribute good outcomes more to luck than a Bayesian would (test of $\gamma_G = 1$: p-value < 0.001), but they are no different from a Bayesian in their treatment of bad outcomes (test of $\gamma_B = 1$: p-value $= 0.267$). Consequently, we find that members tend to attribute good

[2]Results from our first-stage regression suggest that the appointment mechanisms are relevant instruments (F-statistic = 35.23).

outcomes more to luck and bad outcomes more to the DM's decision, and this effect is statistically significant (test of $\gamma_G = \gamma_B$: p-value = 0.042).

Table D.4: IV regression of members' posterior beliefs

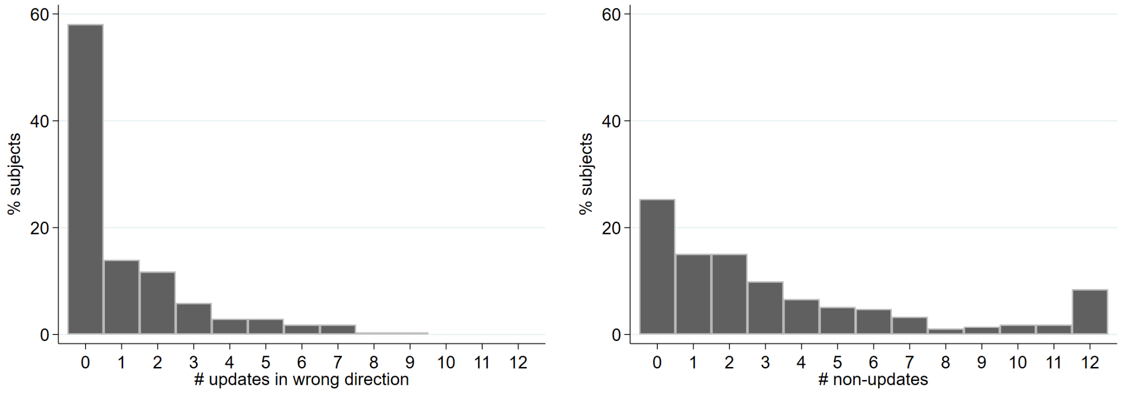
	Dependent variable: Logit(posterior)
	(1)
Variables	Pooled
δ : logit(interim belief)	0.792***
	(0.046)
γ_G : Good outcome × logit(p)	0.787***
	(0.056)
γ_B : Bad outcome × logit($1 - p$)	0.929
	(0.064)
Observations	2,460
# subjects (clusters)	205
Test of $\gamma_G = \gamma_B$	
test statistic	−2.030
p-value	0.042

Robust standard errors clustered at the subject level in parentheses. This analysis excludes subjects classified as inconsistent or non-updaters.
*** p<0.01, ** p<0.05, * p<0.10. Null hypothesis is coefficient = 1.

D.3 Robustness checks of posterior beliefs analysis with inconsistent updaters and non-updaters

This section presents the analysis of members' updating behavior in Experiment 1 with the full sample (i.e., including both inconsistent updaters and non-updaters), as well as with the use of different criteria to exclude inconsistent updaters and non-updaters.

Figure D.1 presents the distribution of subjects based on the number of inconsistent updates and non-updates throughout the experiment. A belief update is classified as *inconsistent* if the posterior belief is in the opposite direction to that predicted by Bayes' rule. A belief update is classified as a *non-update* if the posterior belief is equal to the interim belief. In the main analysis in the paper, we exclude a subject if 25% or more of their posterior beliefs are inconsistent or if they report a posterior belief equal to the interim belief across all six rounds of the experiment.



(a) Inconsistent updates

(b) Non-updates

Figure D.1: Distribution of inconsistent and non-updates by subjects

Table D.5 presents results from the OLS estimation of equation (2) with the inclusion of these previously excluded inconsistent and non-updaters. Overall, we find that the inclusion of these subjects leads to an attenuation of the estimates for γ_G and γ_B. Consequently, at the pooled level (column 1), the estimates now reveal that members tend to attribute bad outcomes more to luck than a Bayesian would (test of $\gamma_B = 1$: p-value = 0.026). This bias is also present at the treatment level, although it is statistically significant in treatments RA, HA, and GA (p-values = 0.056, 0.087, and 0.011, respectively), but not in treatment LA (p-value = 0.379).

Table D.5: Regression of members' posterior beliefs (entire sample)

	Dependent variable: Logit(posterior)				
	(1)	(2)	(3)	(4)	(5)
Variables	Pooled	RA	LA	HA	GA
δ : logit(interim belief)	0.701***	0.737***	0.709***	0.716***	0.539***
	(0.039)	(0.068)	(0.047)	(0.060)	(0.106)
γ_G : Good outcome $\times$ logit(p)	0.530***	0.548***	0.358***	0.618***	0.662***
	(0.064)	(0.086)	(0.094)	(0.083)	(0.093)
γ_B : Bad outcome $\times$ logit($1-p$)	0.848**	0.830*	0.903	0.867*	0.742**
	(0.068)	(0.089)	(0.110)	(0.078)	(0.100)
Observations	3,264	544	1,088	1,088	544
# subjects (clusters)	272	272	272	272	272
R-squared	0.550	0.620	0.606	0.488	0.382
Test of $\gamma_G = \gamma_B$					
test statistic	−3.218	−2.132	−3.376	−2.060	−0.550
p-value	0.001	0.034	0.001	0.040	0.583

Robust standard errors clustered at the subject level in parentheses. This analysis includes all subjects.
*** p<0.01, ** p<0.05, * p<0.10. Null hypothesis is coefficient = 1.

Despite the attenuation in the estimates for γ_G and γ_B, we still find statistically significant evidence that members attribute good and bad outcomes asymmetrically. In particular, even with the inclusion of inconsistent updaters and non-updaters, the estimates in Table D.5 suggest that members tend to attribute good outcomes more to luck and bad outcomes more to the DM's decision (i.e., $\gamma_G < \gamma_B$). This effect is statistically significantly at the pooled level (p-value = 0.001) and in treatments RA, LA, and HA (p-values = 0.034, 0.001, and 0.040, respectively).

We next consider different criteria for excluding non-updaters and inconsistent updaters. Table D.6 presents estimates of members' updating behavior when we exclude rounds where a subject did not make an update for both a good and a bad outcome (column 1), and when we include only subjects who make an update in all six rounds of the task (column 2). Hence, in this table we vary our criteria for excluding subjects based on non-updates. Next, Table D.7 presents estimates of members' updating behavior when we exclude subjects classified as non-updaters and at the same time consider different cut-offs as the criterion for excluding inconsistent updaters (columns 1-9). The maximum number of inconsistent updates by any subject is 9. Hence, column (10) only excludes subjects classified as non-updaters, i.e., those subjects who have not made a single update in all six rounds of the task. Note that column (3) is the criterion used in the paper and corresponds to column (1) of Table 4.

Both tables re-produce the results in Table 4 even when we consider different conditions for excluding non-updates or different thresholds for excluding inconsistent updaters or non-updaters. Importantly, base-rate neglect is always observed (in the first row of both tables). Moreover, in both tables, we systematically observe that good outcomes are attributed to luck (second row), and that good and bad outcomes are treated asymmetrically (last row). Hence, our results are robust to using different criteria for excluding both inconsistent and non-updaters.

Table D.6: Regression of members' posterior beliefs (different criteria for excluding non-updaters)

	Dependent variable: Logit(posterior)	
	(1)	(2)
	Include only subject-rounds	Include only subjects
Variables	with updates	with updates in all rounds
δ : logit(interim belief)	0.465***	0.297***
	(0.047)	(0.131)
γ_G : Good outcome $\times$ logit(p)	0.715***	0.581***
	(0.078)	(0.156)
γ_B : Bad outcome $\times$ logit$(1-p)$	0.957	1.211
	(0.078)	(0.132)
Observations	2,692	954
# subjects (clusters)	249	87
R-squared	0.346	0.286
Test of $\gamma_G = \gamma_B$	-0.242	-0.631
test statistic	-2.010	-2.431
p-value	0.046	0.017

Robust standard errors clustered at the subject level in parentheses. This analysis considers different cut-offs for non-updates. Column (1) includes only subject-rounds with belief updates. Column (2) includes only subjects who have revised their beliefs in all rounds of the investment tasks.
*** p<0.01, ** p<0.05, * p<0.10. Null hypothesis is coefficient = 1.

Table D.7: Regression of members' posterior beliefs (different criteria for excluding inconsistent updaters)

	Dependent variable: Logit(posterior)									
	Criteria for excluding subjects: # inconsistent updates									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	≥ 1	≥ 2	≥ 3	≥ 4	≥ 5	≥ 6	≥ 7	≥ 8	≥ 9	≥ 10
δ : logit(interim belief)	0.783***	0.759***	0.695***	0.667***	0.650***	0.652***	0.653***	0.656***	0.656***	0.656***
	(0.040)	(0.038)	(0.039)	(0.039)	(0.042)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
γ_G : Good outcome $\times$ logit(p)	0.786***	0.728***	0.751***	0.752***	0.742***	0.661***	0.653***	0.607***	0.606***	0.602***
	(0.060)	(0.055)	(0.051)	(0.052)	(0.052)	(0.064)	(0.066)	(0.069)	(0.069)	(0.069)
γ_B : Bad outcome $\times$ logit($1-p$)	1.072	1.030	0.966	0.902	0.873*	0.923	0.913	0.916	0.909	0.905
	(0.088)	(0.074)	(0.067)	(0.068)	(0.067)	(0.072)	(0.072)	(0.072)	(0.072)	(0.072)
Observations	1,620	2,076	2,460	2,652	2,748	2,844	2,904	2,964	2,976	2,988
# subjects (clusters)	135	173	205	221	229	237	242	247	248	249
R-squared	0.710	0.692	0.608	0.565	0.548	0.524	0.518	0.504	0.503	0.503
Test of $\gamma_G = \gamma_B$										
t-statistic	−3.401	−4.216	−3.190	−2.058	−1.771	−2.575	−2.499	−2.862	−2.815	−2.825
p-value	0.001***	0.000***	0.002***	0.041**	0.078*	0.011**	0.013**	0.005***	0.005***	0.005***

Robust standard errors clustered at the subject level in parentheses. This analysis excludes subjects classified as non-updaters (using the strictest criteria) and considers different cut-offs for inconsistent updaters. Column (3) is the benchmark used in the paper. The maximum number of inconsistent updates by any subject is 9. Hence, column (10) only excludes non-updaters.

*** p<0.01, ** p<0.05, * p<0.10. Null hypothesis is coefficient = 1.

D.4 Heterogeneity in updating behavior

Estimates at the pooled level may mask heterogeneity in members' updating behavior. To explore this further, Table D.8 presents the results from both a 2-component (column 1) and 3-component (column 2) finite mixture model analysis of members' updating behavior at the pooled level.[3]

In both models considered in Table D.8, component 1 constitute the majority of updates in the sample (88.9% of the updates in the 2-component and 65.9% of the updates in the 3-component model). This component is characterized by a low level of base-rate neglect and under-responsiveness to both good and bad outcomes. Within this group of belief updates, in relative terms, members attribute good outcomes more to luck and bad outcomes more to the DM's decision, although this difference is statistically significant in the 2-component model (p-value = 0.001) but not in the 3-component model (p-value = 0.190).

The estimates in the table reveal that belief updates in the remaining sample suffer from a higher level of base-rate neglect. Moreover, the under-responsiveness to outcomes is no longer present within this group of updates. Instead, group members are over-responsive to outcomes in component 2 (11.1% of the updates in the 2-component and 4.8% of the updates in the 3-component model). In addition, a third sub-group is identified in the 3-component model (constituting 29.4% of the sample) where members respond to outcomes like a Bayesian.

Overall, our finite mixture model analysis suggests that there is heterogeneity in members' updating behavior. Although members consistently suffer from base-rate neglect, for most updates this is at a modest level. Moreover, the majority of belief updates in the sample is characterized by under-responsiveness to the DM's outcomes and an asymmetric attribution of the DM's outcomes to his/her decision and luck.

[3] We also consider a 4-component model which does not change our main conclusions and does not provide further insight.

Table D.8: Finite mixture model for updating behavior

	Dependent variable: Logit(posterior)	
	2-Component Model	3-Component Model
	(1)	(2)
Component 1		
δ : logit(interim belief)	0.936***	0.972***
	(0.011)	(0.005)
γ_G : Good outcome $\times$ logit(p)	0.535***	0.431***
	(0.031)	(0.031)
γ_B : Bad outcome $\times$ logit($1-p$)	0.668***	0.477***
	(0.044)	(0.040)
Test of $\gamma_G = \gamma_B$		
t-statistic	-3.47	-1.31
p-value	0.001	0.190
Component 2		
δ : logit(interim belief)	0.148***	-0.109***
	(0.086)	(0.137)
γ_G : Good outcome $\times$ logit(p)	1.936**	3.566***
	(0.405)	(0.807)
γ_B : Bad outcome $\times$ logit($1-p$)	1.945**	2.942***
	(0.407)	(0.642)
Test of $\gamma_G = \gamma_B$		
t-statistic	-0.02	0.70
p-value	0.984	0.485
Component 3		
δ : logit(interim belief)		0.302***
		(0.031)
γ_G : Good outcome $\times$ logit(p)		1.029
		(0.054)
γ_B : Bad outcome $\times$ logit($1-p$)		1.103
		(0.067)
Test of $\gamma_G = \gamma_B$		
t-statistic		-0.89
p-value		0.372
Latent Class Marginal Probabilities		
μ_1	0.889	0.659
	(0.020)	(0.028)
μ_2	0.111	0.048
	(0.020)	(0.009)
μ_3		0.294
		(0.027)
Model Fit		
Log likelihood	-3317.86	-3028.11
AIC	6653.720	6084.223
BIC	6705.991	6165.534

Robust standard errors clustered at the subject level in parentheses. This analysis excludes subjects classified as inconsistent or non-updaters.
*** p<0.01, ** p<0.05, * p<0.10. Null hypothesis is coefficient = 1.

D.5 Subjects' behavior in Experiment 2

Table D.9 presents summary statistics of subjects' characteristics in Experiment 1 and Experiment 2. The table reveals significant differences in the subject pool between Experiment 1 and Experiment 2, but not between treatments S and D in Experiment 2 (p-values for test of joint equality are < 0.001 and 0.750, respectively). On average, subjects in Experiment 2 are older (p-value < 0.001), less likely to be Australian (p-value < 0.001) or majoring in economics (p-value $= 0.001$), more likely to be a postgraduate student (p-value < 0.001), and more experienced with economics experiments (p-value < 0.001). However, the subjects do not differ on their decisions in the dictator game or in the risk task between Experiment 1 and Experiment 2 (p-values $= 0.125$ and 0.565, respectively).

Table D.9: Subjects' characteristics in Experiment 1 and Experiment 2

	Exp 2				
	Treatment S	Treatment D	Exp 1	S vs. D	Exp 1 vs. Exp 2
	(1)	(2)	(3)	p-value	p-value
Age	24.963	25.388	21.478	0.379	< 0.001***
	[4.735]	[6.092]	[5.843]		
Female	0.596	0.592	0.537	0.936	0.135
	[0.512]	[0.502]	[0.521]		
Economics major	0.064	0.078	0.140	0.554	0.001***
	[0.245]	[0.268]	[0.347]		
Postgraduate student	0.572	0.515	0.246	0.201	< 0.001***
	[0.496]	[0.501]	[0.432]		
Australian	0.226	0.248	0.971	0.568	< 0.001***
	[0.419]	[0.433]	[0.169]		
# past experiments	2.754	2.874	0.908	0.646	< 0.001***
	[2.380]	[3.458]	[1.757]		
Amount transferred in DG	29.906	30.785	32.809	0.667	0.125
as % of endowment	[21.739]	[23.569]	[21.117]		
Amount invested in RT	71.104	70.680	72.103	0.860	0.565
as % of endowment	[27.175]	[25.414]	[28.220]		
Observations	206	297	272		

Standard deviations in parentheses.
DG: Dictator Game; RT: Risk Task.
*** p<0.01, ** p<0.05, * p<0.10.

We next examine whether members' hypothetical effort choices as DMs in treatment S are consistent with their incentivized decisions in the dictator game. Table D.10 presents marginal-effect estimates from probit regressions of members' hypothetical choices as DMs in round 1 only (column 1) and in all rounds of the task (column 2). We find a statistically significant and positive relationship between members' hypothetical effort choices as DMs and their incentivized decisions in the dictator game (p-values < 0.001 in both columns). A member who transfers 1% more of their endowment to their matched partner in the dictator game is 0.7% more likely to state that they would have chosen high effort as a DM in round 1 of the task. Across all rounds of the investment task, these members are on average 0.5% more likely to state that they would have chosen high effort as DMs.

Table D.10: Regression of members' hypothetical effort choices as DMs in treatment S

	Dependent variable: =1 if member would have chosen e_H as DM	
	Round 1 only	All rounds
Variables	(1)	(2)
% endowment transferred in DG	0.007***	0.005***
	(0.001)	(0.001)
% endowment invested in RT	0.001	0.000
	(0.001)	(0.001)
Game 2	0.043	0.034
	(0.072)	(0.030)
Game 0	0.069	0.010
	(0.072)	(0.034)
Order Effects	N	Y
Observations	198	594
# subjects (clusters)	198	198

Marginal effects of probit model reported. Robust standard errors in parentheses. Standard errors are clustered at the subject level. Game 1 is the reference group in both regressions. DG: Dictator Game; RT: Risk Task.
*** p<0.01, ** p<0.05, * p<0.10.

We next compare members' updating behavior between treatments S and D, and between Experiment 1 and Experiment 2. Table D.11 presents parameter estimates of equation (2) by members' effort choices in treatments S and D, respectively. In the table, we present p-values of pairwise comparisons of the estimates both within each treatment and between treatments. A pairwise comparison of the estimates in columns (3) and (4) reveal that, when subjects play both roles in the experiment, those who choose low effort as DMs are more likely to, as members, attribute good outcomes to luck than those who choose high effort as DMs (p-value = 0.024). This is consistent with members' behavior both in treatment S and in Experiment 1. Moreover, when we examine the updating behavior of subjects separately based on their effort choices as DMs, we do not find any systematic differences between treatments S and D in their attribution of both good and bad outcomes (p-values = 0.264 and 0.190, respectively).[4]

Table D.12 presents parameter estimates of members' updating behavior at the pooled level in treatments S and D (columns 1 and 2). Pairwise comparisons of the estimates reveal that there are no systematic differences in members' updating behavior between the two treatments. Hence, in column (3), we report the parameter estimates by pooling together the subjects in both treatments. The estimates reveal that, overall in Experiment 2, members tend to attribute good outcomes more to luck as compared to a Bayesian (p-value = 0.086). When we compare the members' updating behavior in Experiment 2 with those in Experiment 1 (column 1 of Table 4), we do not find any statistically significant differences in the members' attribution of both good and bad outcomes (p-values = 0.493 and 0.475, respectively), although we find that members in Experiment 2 suffer from stronger base-rate neglect than those in Experiment 1 (p-value = 0.008).

[4]Note that subjects who choose high effort as DMs in treatment D suffer from a stronger base-rate neglect than those who would have chosen high effort hypothetically as DMs in treatment S (p-value = 0.054).

Table D.11: Regression of members' posterior beliefs based on effort choice as DMs in Experiment 2 (treatments S and D)

	Dependent variable: Logit(posterior)							
	Treatment S			Treatment D			S vs. D	
	(1)	(2)	(1) vs. (2)	(3)	(4)	(3) vs. (4)	(1) vs. (3)	(2) vs. (4)
Variables	Low effort	High effort	p-value	Low effort	High effort	p-value	p-value	p-value
δ : logit(prior belief)	0.483***	0.607***	0.254	0.571***	0.339***	0.098*	0.440	0.054*
	(0.079)	(0.077)		(0.083)	(0.115)			
γ_G : Good outcome × logit(p)	0.601*	1.068	0.085*	0.730	1.383*	0.024**	0.656	0.264
	(0.210)	(0.173)		(0.201)	(0.223)			
γ_B : Bad outcome × logit($1-p$)	1.181	0.759	0.082*	0.989	1.074	0.734	0.459	0.190
	(0.188)	(0.171)		(0.179)	(0.169)			
Observations	656	232		694	320			
# subjects (clusters)	137	71		147	90			
R-squared	0.392	0.488		0.432	0.323			

Robust standard errors clustered at the subject level in parentheses. This analysis excludes subjects classified as inconsistent or non-updaters.
Columns (1) and (2) control for members' hypothetical effort choices in treatment S, and are the same as columns (3) and (4) of Table 6.
*** p<0.01, ** p<0.05, * p<0.10. Null hypothesis in columns (1)-(4) is coefficient = 1.

Table D.12: Regression of members' posterior beliefs at pooled level in Experiment 2 (treatments S and D)

	Dependent variable: Logit(posterior)				
	Treatment S	Treatment D	S + D	(1) vs. (2)	(3) vs. Exp 1[a]
Variables	(1)	(2)	(3)	p-value	p-value
δ : logit(prior belief)	0.525***	0.546***	0.535***	0.815	0.008***
	(0.063)	(0.064)	(0.045)		
γ_G : Good outcome $\times$ logit(p)	0.803	0.845	0.827*	0.836	0.493
	(0.146)	(0.139)	(0.100)		
γ_B : Bad outcome $\times$ logit($1-p$)	0.990	1.107	1.052	0.537	0.475
	(0.145)	(0.122)	(0.094)		
Observations	888	1,014	1,902		
# subjects (clusters)	148	169	317		
R-squared	0.399	0.400	0.399		

Robust standard errors clustered at the subject level in parentheses. This analysis excludes subjects classified as inconsistent or non-updaters.

[(a)] This refers to the estimates reported in column (1) of Table 4.

*** p<0.01, ** p<0.05, * p<0.10. Null hypothesis in columns (1)-(3) is coefficient = 1.

E Detailed explanation of estimation strategy for posterior beliefs

In this section, we explain in detail the interpretations of the parameters presented in Section 2.3.3.

We first express posterior beliefs in terms of log likelihood ratios. We have

$$\log\left(\frac{\phi_i^\Psi|_{Q_H}}{1-\phi_i^\Psi|_{Q_H}}\right) = \log\left(\frac{\mu_i^\Psi}{1-\mu_i^\Psi}\right) + \log\left(\frac{p}{1-p}\right), \tag{E.1}$$

and

$$\log\left(\frac{\phi_i^\Psi|_{Q_L}}{1-\phi_i^\Psi|_{Q_L}}\right) = \log\left(\frac{\mu_i^\Psi}{1-\mu_i^\Psi}\right) + \log\left(\frac{1-p}{p}\right). \tag{E.2}$$

By letting $\text{logit}(x) \equiv \log\left(\frac{x}{1-x}\right)$, we can jointly express (E.1) and (E.2) as

$$\text{logit}(\phi_i^\Psi|_Q) = \text{logit}(\mu_i^\Psi) + I(Q = Q_H) \cdot \text{logit}(p) + I(Q = Q_L) \cdot \text{logit}(1-p), \tag{E.3}$$

where $I(\cdot)$ is an indicator function.

Equation (2) in the paper is obtained by augmenting equation (E.3) in the following way:

$$\text{logit}(\hat{\phi_i^\Psi}|_Q) = \delta\,\text{logit}(\hat{\mu_i^\Psi}) + \gamma_G\, I(Q = Q_H) \cdot \text{logit}(p) + \gamma_B\, I(Q = Q_L) \cdot \text{logit}(1-p) + \varepsilon_i, \tag{E.4}$$

where ε_i captures non-systematic errors. This specification allows us to determine the weights members place on their interim beliefs and the signals they receive. Note that $\delta = \gamma_G = \gamma_B = 1$ equates (E.4) to (E.3). This is the case where there is no bias in belief updating.

Any deviation in the parameters from 1 is interpreted as non-Bayesian updating behavior. Specifically, δ captures the weight that a group member places on his/her interim belief in the updating process, γ_G captures the extent to which a member responds to a signal of good outcome from the DM, and γ_B captures the extent to which a member responds to a signal of bad outcome from the DM. We use Figures E.1 and E.2 to explain these parameters in more detail.

Figure E.1 shows the implications of different values of δ on the relationship between the member's posterior and interim beliefs, conditional on observing a good outcome and holding γ_G constant (at 1).[5] Note that δ corresponds to the slope of the linear regression. If $\delta < 1$, then the member suffers from *base-rate neglect* in that s/he places too little weight on his/her interim belief. To see this, consider a member whose interim belief μ_A is less than 0.5. This corresponds to $\text{logit}(\mu_A) < 0$ in Figure E.1. Hence, the member believes that the DM is more likely to have chosen low effort. When Q_H is observed, the signal contradicts with the interim belief. However, s/he arrives at a posterior belief that

[5] A similar analysis can be done for the case where a bad outcome is observed.

is greater than that of a Bayesian (i.e., point A' instead of point A). In other words, the member neglects his/her interim belief and over-updates in response to receiving a signal that contradicts with what s/he initially believes to be true.[6]

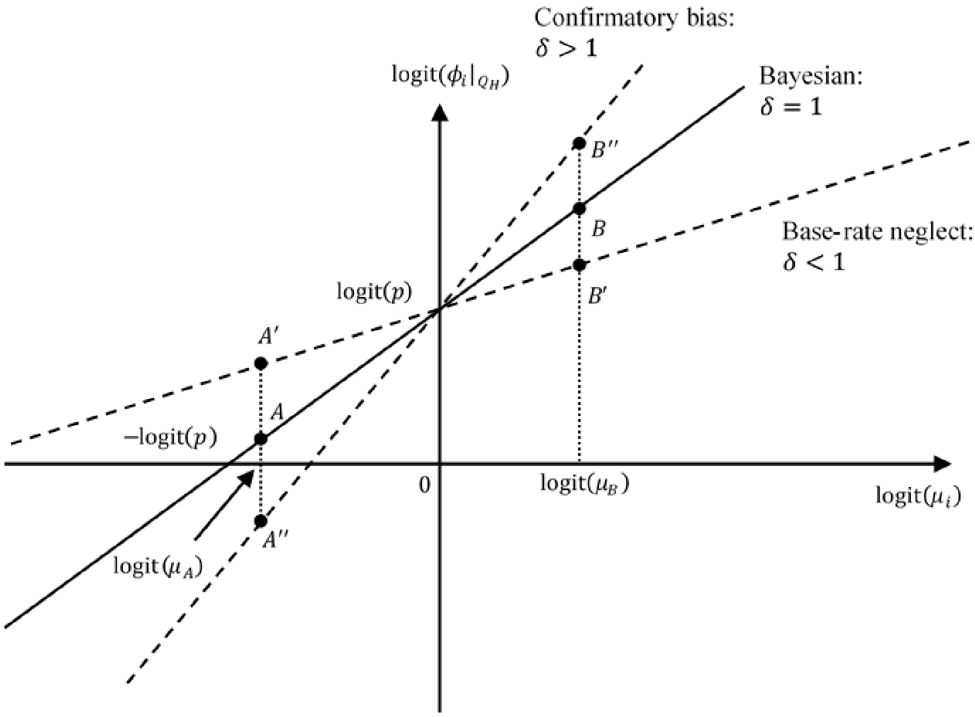


Figure E.1: Interpretation of δ given Q_H observed and $\gamma_G = 1$

Conversely, $\delta > 1$ implies that the member suffers from *confirmatory bias* in that s/he places too much weight on his/her interim belief. To see this, consider a member whose interim belief μ_B is greater than 0.5, i.e., $\text{logit}(\mu_B) > 0$ in Figure E.1. When Q_H is observed, the signal confirms the interim belief. However, his/her posterior belief is at point B'' instead of point B. Hence, the member over-updates relative to a Bayesian when s/he receives a signal that confirms what s/he initially believes to be true.[7]

Figure E.2 shows the implications of different values of γ_G on the relationship between the member's posterior and interim beliefs.[8] Note that γ_G and γ_B correspond to the intercepts of the regression conditional on the signal received by the member. If $\gamma_G > 1$, then the member is, on average, over-responsive to good signals relative to a Bayesian, and tends to arrive at a posterior that is higher than that of a Bayesian. Specifically,

[6]Now consider a member whose interim belief μ_B is greater than 0.5. After observing Q_H, a signal that confirms this belief, suppose that his/her posterior belief is at B'. This implies that a member who suffers from base-rate neglect does not update as much as a Bayesian would when s/he receives a signal that confirms his/her interim belief.

[7]Alternatively, consider a member whose interim belief μ_A is less than 0.5. After observing Q_H, a signal that contradicts with this belief, suppose that his/her posterior belief is at A''. This implies that a member who suffers from confirmatory bias does not update as much as a Bayesian would when s/he receives information that contradicts with his/her interim belief.

[8]A similar analysis can be done for the case where a bad outcome is observed.

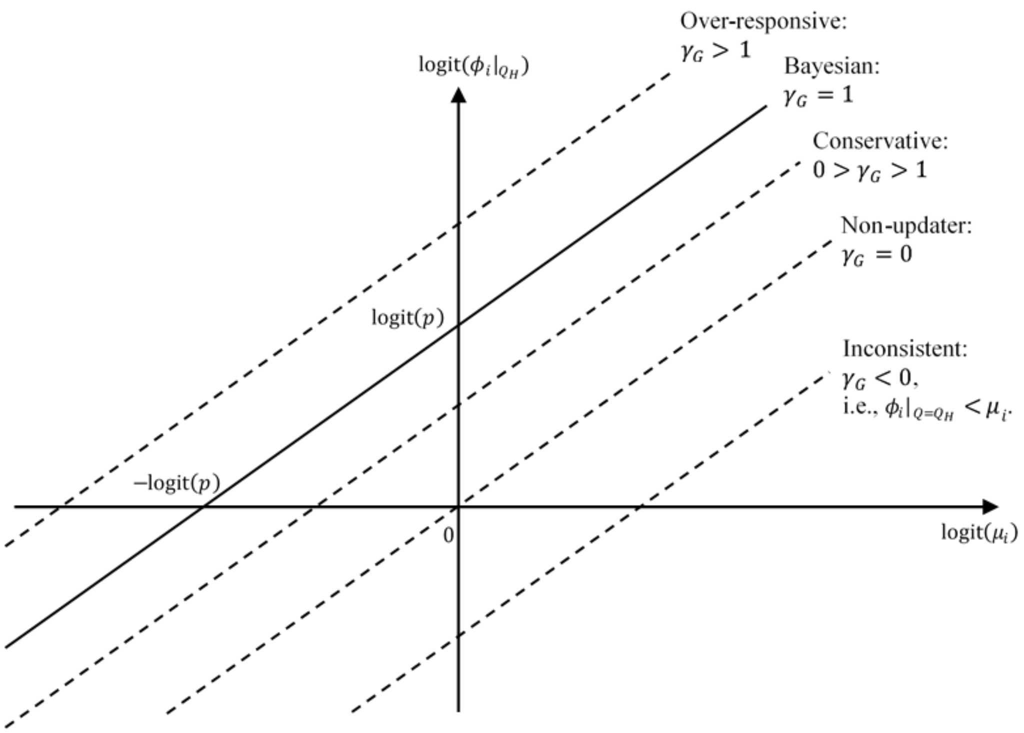


Figure E.2: Interpretation of γ_G given Q_H observed and $\delta = 1$

the biased member attributes good outcomes more to the DM's decision as compared to an unbiased Bayesian member. On the other hand, if $\gamma_G < 1$, then the member is conservative in his/her response to good signals, and tends to arrive at a posterior that is lower than that of a Bayesian on average. In this case, the biased member attributes good outcomes more to luck as compared to an unbiased Bayesian member. Figure E.2 also shows what happens when $\gamma_G = 0$ or $\gamma_G < 0$, which correspond to a non-updater and an inconsistent updater, respectively.

Finally, we can also capture asymmetric updating of beliefs, i.e., asymmetric attribution of outcomes to the DM's decision (effort choice) and luck. If $\gamma_G > \gamma_B$, then the member is more likely to attribute a good outcome to the DM's decision and a bad outcome to luck. Conversely, if $\gamma_G < \gamma_B$, then the member is more likely to attribute a bad outcome to the DM's decision and a good outcome to luck.