Journal of Experimental Social Psychology 92 (2021) 104067

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Social deliberation systematically shifts resource allocation decisions by focusing on the fate of the least well-off

Atsushi Ueshima[a,b], Hugo Mercier[c], Tatsuya Kameda[a,d,*]

[a] Department of Social Psychology, The University of Tokyo, Tokyo, Japan
[b] Japan Society for the Promotion of Science, Tokyo, Japan
[c] Institut Jean Nicod, Département d'études cognitives, ENS, EHESS, PSL University, CNRS, Paris, France
[d] Center for Experimental Research in Social Sciences, Hokkaido University, Hokkaido, Japan

ARTICLE INFO

Keywords:
Resource allocation
Deliberation
Argumentation
Maximin
Equality

ABSTRACT

How much inequality should be tolerated? How should the poorest be treated? Though sometimes conflated, concerns about inequality and the fate of the poorest involve different allocation principles with different sociopolitical implications. We tested whether deliberation—the core of democracy—influences reasoning about distributive principles. 322 participants faced allocation decisions for others between egalitarian (low variance in allocation), utilitarian (high total amount), and maximin (maximizing the welfare of the poorest) options. After their initial decisions, participants either reflected upon similar decisions solely or discussed them in pairs before facing the same choices again individually. Social, but not solitary, deliberation led to more maximin and fewer egalitarian choices, and this change lasted at least 5 months after the experiment. Conversation analyses of approximately 7500 utterances suggest that some participants initially made egalitarian choices heuristically, when in fact they mostly cared about the poorest, and dialogue promoted more internally coherent maximin preferences.

1. Introduction

Debates over what constitutes a fair distribution of resources, how much inequality a society should tolerate, or how less well-off citizens should be treated have always been central to democratic societies. The public's preferences over resource allocations are of crucial practical relevance, as public opinion shapes policy (e.g., Burstein, 2003). These preferences are also important from a theoretical point of view, as they relate to theories of human morality: Are humans broadly adverse to inequality (Dawes, Fowler, Johnson, McElreath, & Smirnov, 2007; Fehr & Schmidt, 1999; Hsu, Anen, & Quartz, 2008)? Or do they prefer fair distributions, even if they entail a degree of inequality (Baumard, André, & Sperber, 2013; Kameda et al., 2016; Starmans, Sheskin, & Bloom, 2017)?

Some studies suggest that humans prefer more egalitarian outcomes. In various economic games (e.g., dictator, ultimatum, and trust games), participants have been shown to be sensitive to inequality and motivated to achieve more equal distributions (see Fehr & Schmidt, 2006 for a review). Likewise, Shaw (2013) has demonstrated that, when individual contributions are equal, people from a young age prefer more equitable resource allocations, even when pursuing equality entails a waste of resources.

However, recent studies have questioned whether many people have a broad preference for egalitarian outcomes (Starmans et al., 2017), or whether in fact their egalitarianism is an indirect way of protecting the less well-off (Kameda et al., 2016). It has been suggested that many apparently egalitarian people care about a Rawlsian *maxi*min principle, attempting to maximize (maxi) the resources of the less well-off (min) (Rawls, 1971). Engelmann and Strobel (2004) have demonstrated that preference for equality is stronger when it aligns with maximin concern, i.e., when restoring equality also means maximizing the welfare of the poorest recipients (see also Charness & Rabin, 2002; Frohlich, Oppenheimer, & Eavey, 1987; Mitchell, Tetlock, Mellers, & Ordóñez, 1993). Kameda et al. (2016) gave participants choices between three distributions of resources: one that maximized total payoff (*utilitarian option*), one that maximized equality (*egalitarian option*), and one that maximized payoff to the worst-off recipient (*maximin option*). Participants directed most of their attention to information about the worst-off recipients, suggesting that, for many of them, those payoffs played an outsized role in their decisions (see also Frohlich & Oppenheimer, 1992).

Whatever participants' initial preferences are, these preferences may be affected by deliberation. In the Cartesian tradition, deliberation is understood as a process of solitary thought and effortful reflection upon one's

* Corresponding author at: Department of Social Psychology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.
E-mail address: tkameda@l.u-tokyo.ac.jp (T. Kameda).

https://doi.org/10.1016/j.jesp.2020.104067
Received 19 May 2020; Received in revised form 27 September 2020; Accepted 29 September 2020

own ideas (Descartes, 1637/2006). Even if such solitary deliberation is doubtlessly productive in some cases, studies suggest that its main effect is often to bolster pre-existing beliefs rather than change them. In many domains, participants engaging in solitary deliberation have been shown to mostly find arguments and evidence that support their initial views, a phenomenon known as confirmation bias or myside bias (for reviews, see Mercier, 2016a; Nickerson, 1998). Because of this myside bias, participants are unlikely to find reasons to call their initial beliefs into question; instead, the piling up of supportive reasons leads to overconfidence (Koriat, Lichtenstein, & Fischhoff, 1980) or polarization (Tesser, 1978).

In contrast to solitary deliberation, social deliberation—exchanging arguments with others (Mercier & Landemore, 2012)—has been repeatedly shown to result in changes of mind. In a variety of domains, when people discuss in small groups, they change their initial judgments or beliefs, for instance by adopting the best answer present in the group when there are normatively correct answers (for reviews, see Laughlin, 2011; Mercier, 2016b; Mercier & Sperber, 2017). Social deliberation is particularly effective when the group members share a frame of reference and incentives but disagree on some issue. In the absence of disagreement, social deliberation often turns into an expanded form of solitary deliberation, with the discussants piling up arguments for the agreed upon opinion (Sunstein, 2002, although there are exceptions to this pattern; see, e.g. Himmelroos & Christensen, 2014).

When group members disagree about the answer to a problem with a normatively correct answer, the contrast between internal and social deliberation is stark. When participants are asked to reflect on their own about a tricky reasoning problem, few shift from an incorrect answer to the correct answer (see, e.g., Bago & De Neys, 2019). By contrast, when participants tackle the same problem in discussion groups, if at least one group member defends the correct answer, then everyone ends up adopting it (Claidière, Trouche, & Mercier, 2017; Laughlin & Ellis, 1986; Tindale, Smith, Thomas, Filkins, & Sheffey, 1996; Trouche, Sander, & Mercier, 2014).

When a group member convinces others to accept the correct answer to a reasoning problem, they do so by showing how the correct answer is more coherent with the agreed upon axioms of the problem (see, e.g., Moshman & Geil, 1998). As a result, participants who change their minds genuinely understand the reasoning behind the correct answer, such that they are able to recreate the argument to convince others in turn (Claidière et al., 2017), and to apply the logic of the argument to transfer problems (Trouche et al., 2014; for review, see Laughlin, 2011).

Social deliberation has also been found to affect decisions in two domains related to preferences over resource allocation. First, experiments with economic games show that groups make more rational decisions than individuals, which can translate into more utilitarian decisions (Kugler, Kausel, & Kocher, 2012). Second, when groups are asked to make risky choices together, a “risky shift” is often observed (Kerr, MacCoun, & Kramer, 1996). This risky shift might translate into resource allocations that maximize total payoffs at the cost of equality or outcomes for the worst-off recipients (i.e. more utilitarian decisions).

However, the conclusion that social deliberation might yield more utilitarian resource-allocation preferences must be qualified. First, the risky shift is sometimes reversed (Masclet, Colombier, Denant-Boemont, & Lohéac, 2009). Second, when participants had to discuss distributional principles in small groups, few groups agreed on the utilitarian outcome (Frohlich et al., 1987). More importantly, people make different decisions when deciding for themselves (as in the experiments mentioned above, including Frohlich et al., 1987) than when deciding for others (e.g. Ogawa, Ueshima, Inukai, & Kameda, 2018; Stone, Yates, & Caruthers, 2002). Of particular relevance, Kameda et al. (2016) showed that participants made more maximin decisions when choosing resource allocations for others than when choosing for themselves. As a result, we cannot draw solid hypotheses about the effect of social deliberation on utilitarian decisions.

By contrast, we can derive hypotheses pertaining to the relationship between equality and maximin outcomes. Although resource allocation decisions do not have a normatively correct answer, social deliberation might play the same role in resource allocation decisions as it does in reasoning problems: making the participants' opinions more internally coherent. As mentioned above, it has been suggested that the concern for equality demonstrated in many studies in fact reflects a preference for the maximin strategy: some participants use maximizing equality in a heuristic manner, as maximizing equality often also increases the payoff for the least well-off resource recipients (Kameda et al., 2016). As a result, when addressing participants who have chosen the most equal option, but who only did so as a heuristic when their real preference is for maximin, other participants should be able to point out that the maximin choice is more coherent with these participants' preferences than the egalitarian choice. From these considerations, we draw the following hypotheses:

H1. : Compared to solitary deliberation, social deliberation will prompt participants to make more maximin decisions.

H2. : Compared to solitary deliberation, social deliberation will prompt participants to make fewer egalitarian decisions.

Past results have shown broad support, after social deliberation, for distributional principles relying largely on maximin, by contrast with utilitarian principles (Frohlich et al., 1987), but these experiments have not compared maximin to egalitarian principles. Moreover, the processes by which groups reached their decision was not studied. In the present experiments, several steps were taken to show that a potential shift from egalitarian to maximin decisions was due to the exchange of arguments (Jensen, 2016), and not to extraneous factors. First, the discussions were between pairs of participants instead of larger groups. In a pair, a participant cannot merely yield to the majority, as many participants do in group decision making (Hastie & Kameda, 2005; Kameda, Tsukasaki, Hastie, & Berg, 2011; Kerr et al., 1996). Still, even in a pair, participants might be influenced by factors besides argument quality, such as confidence (Bahrami et al., 2010; Trouche et al., 2014). To further ensure that arguments play an important role, we took two more steps. We examined the content of the discussions to test the following hypothesis:

H3. : Mentions of the worst-off recipients during discussion relate to shifts from an egalitarian to a maximin strategy.

We also examined whether the potential shift towards maximin options applied to decisions besides those discussed, which allowed us to test the following hypothesis:

H4. : H1 and H2 are observed for allocation choice problems that have not been discussed but have the same structure with different numbers.

In the first phase of Study 1, participants made a series of decisions between egalitarian, utilitarian, and maximin options. In the second phase, they made several more similar decisions either on their own (Solo Condition) or paired with another participant with the instructions to reach a consensus (Pair Condition). Finally, in the third phase, they were confronted with the same series of decisions as the first phase. We compare the evolution of decisions from the first to the third phase between the Solo and Pair conditions. Study 2 replicates Study 1 while improving on several aspects (e.g., superior process measurement during the second phase). Finally, we conducted a follow-up online study that was planned in advance along with Study 2, in which participants from Study 2 were confronted again with the same decisions as in the first and third phases, but five months after Study 2 took place.

2. Study 1

2.1. Method

2.1.1. Participants

A total of 156 student volunteers at a Japanese university participated in the experiment (99 male; M_{age} = 19.3, SD = 0.95). Minimum sample size was determined prior to the start of data collection as

follows. In conversation analysis using a repeated measures ANOVA with three within-factors (Low, Medium, and High) given $\alpha = 0.05$ and power = 0.80, a sample size of 50 pairs (100 participants) allows us to detect an effect as small as 0.18 (a small-to-medium effect using Cohen's *f*). Although we planned to use a mixed-effects Poisson regression for conversation analysis in advance, here we based the sample size calculation on a repeated measures ANOVA instead, because a practical method for conducting sensitivity analysis with mixed-effects Poisson regression is not yet well-established, to the best of our knowledge. Also, we focused on conversation analysis to determine the sample size, as the conversation analysis would involve the smallest number of analysis units (50 pairs) in our experiments. This led us to have at least 50 pairs in the Pair condition, and accordingly at least 50 individuals in the Solo condition (conditions to be explained below). This yielded a total of at least 150 participants. No data analysis was conducted before we finished data collection. Informed consent was obtained from each participant using a consent form approved by the Institutional Review Board of the Center for Experimental Research in Social Sciences at the University.

2.1.2. Choice problem

Participants were provided a choice set composed of three options in each trial (Fig. 1A): one with the largest total ("Utilitarian" option), one with the smallest variance in terms of the Gini coefficient ("Egalitarian" option), and one with the largest minimum ("Maximin" option). Participants were asked to choose one of these options as a third-party allocation to three others, and were told that they and the recipients would remain mutually anonymous. We used the same 40 choice sets as those used in a previous study (Kameda et al., 2016) (see Table S1 for the full list of the choice sets and information about their structure). It is important to note that we predicted that the experience of pair discussion would have a specific influence on maximin and egalitarian preferences—increase in maximin preferences and decrease in egalitarian preferences (H1 and H2). We did not predict that maximin preferences would increase at the expense of social efficiency (i.e., utilitarian preferences). In order to test this point, we included options that corresponded to utilitarian preferences as well as maximin and egalitarian preferences in our distribution task.

2.1.3. Procedure

A schematic illustration of the decision settings is shown in Fig. 1B. Upon arrival, participants were seated in private cubicles and read an explanation of the entire experimental session. They were informed that: (a) the entire experimental session would consist of three phases in which they would make distributive choices for unknown others who were participating in a different experiment, (b) one allocation choice from the entire session would later be selected randomly to determine the recipients' actual monetary outcomes to be paid in cash, (c) participants and the recipients would remain mutually anonymous. To be consistent with previous distributive studies (e.g., Engelmann & Strobel, 2004; Kameda et al., 2016; Ogawa et al., 2018), we made the allocation task abstract and did not provide specific details about the recipients or the experiment they were supposed to participate in.

In the first and the third phases, our participants made 40 allocation choices to three unknown others (labeled persons A, B, and C in the first phase, and persons H, I, and J in the third phase). The presentation order of 40 problems and the locations of the three options (Egalitarian, Maximin, or Utilitarian) on the screen were randomized across trials. In each trial, participants confirmed their choices by clicking a check mark at the left of the chosen option. Participants were asked to make choices within 15 s for each problem. Before making actual choices, participants answered a quiz testing their understanding of the task. The first phase was identical for participants in the Solo and Pair conditions, and those in the Pair condition were not informed in advance that they would discuss the distribution problems with someone else in the second phase. Stimuli were presented on a laptop computer (PC-GN256FSG8, NEC co., Japan) using a PsychoPy script (Peirce et al., 2019).

In the second phase, the participants were asked to make five allocation choices again to three anonymous recipients (labeled persons

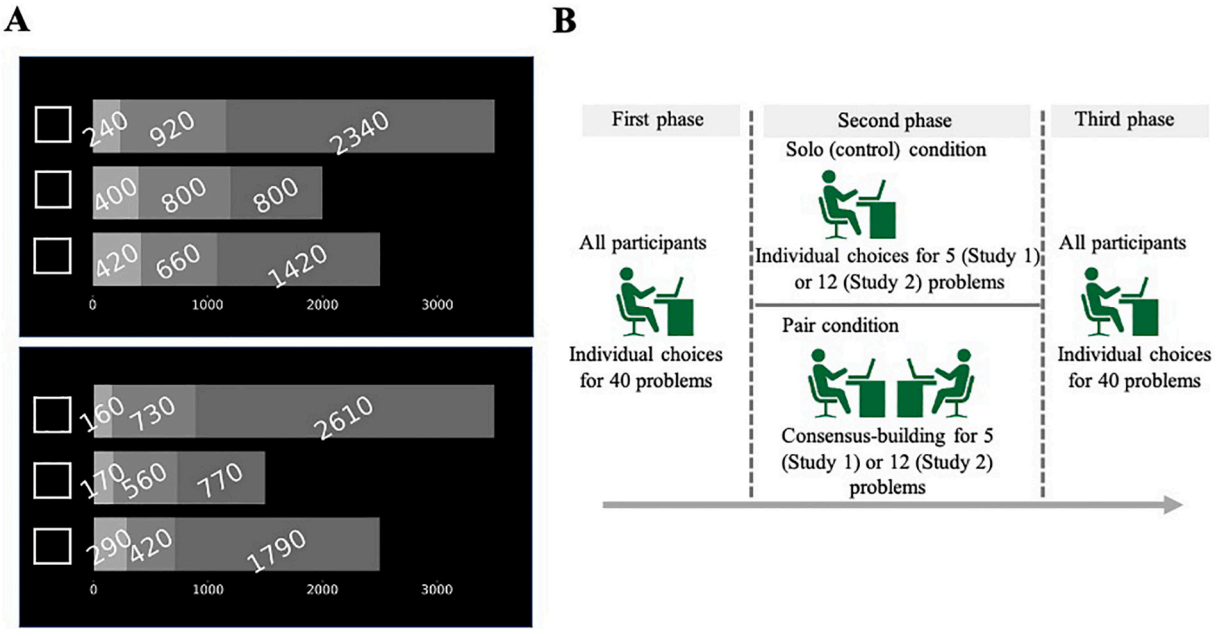


Fig. 1. A schematic illustration of the task design. (A) Two examples of a choice set composed of three options (in Yen). See Table S1 for a full list of the 40 choice sets used in the experiments. In these examples, the top option represents a Utilitarian allocation, the middle option an Egalitarian allocation, and the bottom option a Maximin allocation. The locations of the three options on the screen were randomized across trials, and the display orders of the amounts (Low-Medium-High or High-Medium-Low, from left to right) were randomized across participants in the Solo condition and across pairs in the Pair condition. (B) Three phases of the experiment. In the second phase, participants deliberated about five pre-selected (Study 1) or 12 randomly-selected (Study 2) problems from the 40 choice sets, either individually or as a pair.

X, Y, and Z). These five allocation problems had been preselected randomly from the 40 choice sets before the experiment and were commonly administered to the participants (choice sets 1, 9, 30, 33, and 39 in Table S1 were used). Participants in the Pair condition ($N = 100$: 24 male-male pairs, 14 female-female pairs, and 12 mixed-gender pairs) were asked to make a consensus decision on each of the five problems. They moved to an adjacent room for discussion. Participants in the Solo condition ($N = 56$, 39 males) worked alone on the same five problems. Each pair or participant was provided with five answer sheets and asked to answer the five allocation problems within 10 min (which was longer than the decision time of 15 s per problem in the first phase). In the Pair condition, we recorded each pair's discussion. Transcription of audio data was conducted by a third party company (Tokyo Hanyaku Co., Ltd.), whose personnel did not know the hypotheses of our study. To segment the transcript texts into words, we used MeCab v. 0.996 (Kudo, 2018) and RMeCab v. 0.99999 (Ishida, 2018) as a Japanese morphological analyzer.

In all three phases, the display orders of the amounts (Low-Medium-High or High-Medium-Low, from left to right; see Fig. 1A) were randomized across participants in the Solo condition and across pairs in the Pair condition. Participants who saw the order of "Low-Medium-High" (see Fig. 1A) were informed that person A (X in the second phase, H in the third phase) would receive the lowest outcome; B (Y, I) the middle outcome; and C (Z, J), the highest outcome. Participants who saw the order of "High-Medium-Low" were informed that person A (X, H) would receive the highest outcome; B (Y, I) the middle outcome; and C (Z, J), the lowest outcome.

Participants received 1000 yen (approximately US$10) for participation at the end of the experiment, and were dismissed. The recipients were also paid later according to the aforementioned procedure.

2.1.4. Data analysis and software

In the analysis, we used Markov chain Monte Carlo (MCMC) methods for parameter estimation. Models were implemented with weakly or non-informative priors using rstan version 2.17.3 (Carpenter et al., 2017) and brms (Bürkner, 2017) with R version 3.4.3 (R Core Team, 2017). Multinomial tests, *t*-tests, and contingency table analysis were implemented with JASP 0.12.2 (JASP Team, 2020). In all analyses other than the analysis using multiple imputation approach, the statistics were below 1.1. Figures showing results were created using ggplot2 (Wickham, 2016). Effect sizes are reported as standardized regression coefficients for regression models with continuous explanatory variables. In all studies, we report all measures, manipulations and exclusions.

2.2. Results

2.2.1. Behavioral data

We analyzed the decisions from the second phase, in which participants were asked to deliberate about five problems individually or as a pair (see Table S1 for the problems used). Fig. 2A displays decision rates for the Maximin, Egalitarian, and Utilitarian options. A multinomial logistic regression analysis (with the Maximin options as a reference and with varying intercepts for participants or pairs) revealed that the Maximin options were chosen more frequently than the Egalitarian options ($\beta = -1.94$, 95% credible interval [−2.82, −1.23]) or the Utilitarian options ($\beta = -1.06$, 95% CI [−2.11, −0.04]). The difference between the Solo and the Pair conditions was not significant ($\beta = -0.07$, 95% CI [−1.03, 0.92] for the Egalitarian options; $\beta = -1.16$, 95% CI [−2.76, 0.29] for the Utilitarian options).

To shed light on the deliberation processes in pairs, Fig. 2B presents a 'social decision scheme matrix' (Davis, 1973; Kameda, Tindale, & Davis, 2003) depicting how each pair aggregated initial individual preferences into pair decisions about the five problems that were repeated across both phases. This matrix summarizes observed conditional probabilities about how the six possible preference-configurations in the first phase were resolved into pair-level decisions in the second phase. It is noteworthy that the Egalitarian-Utilitarian pairs (see the bottom row), in which neither member had initially endorsed the Maximin option, chose the Maximin option much more frequently (0.67) as compared to the two options according to their original preferences (0.07, 0.27 respectively). This pattern was not statistically predictable from a model assuming that discussion would yield only either Egalitarian or Utilitarian consensus with equal (i.e., 0.5) probability (Bayes factor > 100; multinomial test), suggesting that the Maximin option emerged as a new viable option through pair dialogue.

To test H1 and H2, which predict that the exchange of arguments yields an overall shift from egalitarian to maximin preferences, we examined whether participants in the pair condition chose the Maximin options more frequently and the Egalitarian options less frequently in the individual decisions from the third phase compared to the first phase (see Fig. S2 for the within-person choice stability of participants—the degree to which each participant chose the same type of options consistently—during the first phase). As hypothesized (see Fig. 3A), participants in the Pair condition made fewer Egalitarian and more Maximin choices compared to those in the Solo condition ($\beta_{Pair*third\ phase} = 0.47$, 95% CI [0.20, 0.74] by a multinomial logistic regression analysis with the Egalitarian options as a reference and with varying intercepts for participants). Importantly, this pattern held when we removed the five choice problems used in the second phase from the analysis ($\beta_{Pair*third\ phase} = 0.41$, 95% CI [0.21, 0.71]), indicating that participants learned the difference between Maximin and Egalitarian logics as transferable knowledge (Gick & Holyoak, 1983), and applied this knowledge to the choice problems that did not appear in the second phase (H4). This pattern was also significant when we included participant gender as a covariate ($\beta_{Pair*third\ phase} = 0.47$, 95% CI [0.20, 0.74]). These results support H1, H2, and H4.

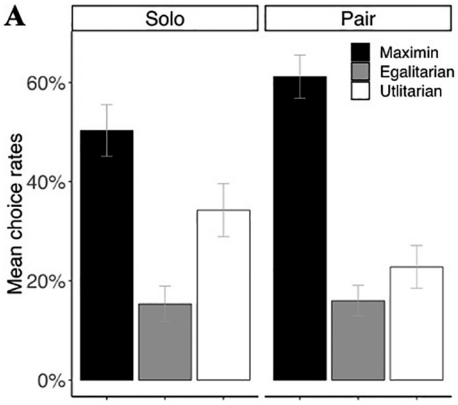
2.2.2. Conversation data

We confirmed that the maximin concern was the key topic during discussion by analyzing how many times each amount (Low, Medium, or High) was mentioned by each pair during discussion (Fig. 3B). The average number of words (nouns, verbs, adjectives, and adverbs) uttered during pair discussion in each group was 251.72 words. Pairs mentioned Low amounts more frequently than Medium or High amounts ($\beta_{medium} = -3.41$, 95% CI [−4.84, −2.26]; $\beta_{high} = -1.71$, 95% CI [−2.31, −1.14] by Poisson regression with varying intercepts for pairs). It is noteworthy that participants generally mentioned the amounts received by the worst-off recipient most often, even though there were substantive variations in their prior individual preferences (Egalitarian and Utilitarian choices accounted for 44% of all choices in the first phase), indicating that the maximin concern operated robustly as a cognitive anchor in pair discussion (Frohlich & Oppenheimer, 1992; Kameda et al., 2016).

To test whether mentions of the worst-off amounts in the discussion related to shifts from an egalitarian to a maximin strategy (H3), we investigated the relation between the frequency of mentions of the minimum amounts and the increase in the number of Maximin choices after pair discussion. As seen in Fig. 3C, however, this relation was not credible ($\beta_{frequency\ low} = 0.34$; 95% CI [−1.42, 2.12]), suggesting that simply mentioning the amounts attributed to the worst-off was not enough to clarify the difference between the maximin and the egalitarian logics. Thus, H3 was not supported.

2.3. Discussion

Study 1 confirmed H1, H2, and H4. Even though the decision patterns in the second phase were not distinguishable between the Solo and the Pair conditions, the pair discussion had the effect of increasing the concern for the poorest, while decreasing the concern for economic inequality among recipients per se in the third phase (H1 and H2). Importantly, this shift was observed not only for the problems seen in the second phase but also for the other problems (H4).



B

Pair decisions

Individual configurations	Maximin	Egalitarian	Utilitarian	Freq.
(Maximin, Maximin)	0.83	0.09	0.08	76
(Egalitarian, Egalitarian)	0.27	0.73	0	11
(Utilitarian, Utilitarian)	0.17	0.06	0.78	18
(Maximin, Utilitarian)	0.51	0.04	0.45	71
(Maximin, Egalitarian)	0.64	0.34	0.02	59
(Egalitarian, Utilitarian)	0.67	0.07	0.27	15

Fig. 2. Choice results in the second (deliberation) phase of Study 1. (A) Mean choice rates of the three options in the Solo and Pair conditions. Error bars represent SEM. (B) Social decision scheme matrix indicating how each pair aggregated initial individual preferences (first phase) into pair decisions (second phase) on the five problems that were used both in the first phase and the pair discussion. The leftmost column indicates the six possible preference configurations in the first phase, and the rightmost column indicates their observed frequencies. Notice that the individual configurations were defined for each of the five problems and each of the 50 pairs, so that the total frequency is 250. For example, the individual configuration of Maximin-Maximin (top row) corresponds to the 76 observed cases in which both participants in a pair had individually chosen the Maximin option on a problem in the first phase. The row entries of the matrix represent observed conditional probabilities that each configuration yielded the Maximin, Egalitarian or Utilitarian decision, respectively, as a pair. See also Fig. S1 for an analysis using individual configurations based on participant's "behavioral type" (Egalitarian, Maximin, or Utilitarian type according to their most-frequent choices for the 40 problems in the first phase), instead of the choice-based configuration for each problem.

However, the mere mention of the minimum amounts during discussion did not relate to the increase in maximin choices after discussion (H3). Previous research has shown that making comparisons between two concepts help in learning their relational structure (Edwards, Williams, Gentner, & Lombrozo, 2019; Gentner, 1983). Thus, to better grasp the maximin logic, participants may need to directly compare the information indicating the maximin option (i.e. payoff for the least well-off) and the egalitarian option (i.e. variance between the payoffs). However, such a comparison was not directly available in Study 1, as no single piece of information precisely conveyed the variance between the payoffs. In Study 2, such summary information was directly provided. Also as seen in Fig. 1A, in Study 1 participants could identify the maximin option by simply comparing the lengths of the leftmost bar segments; that is, in Study 1, participants could rely on the options' visual features, instead of examining the precise numerical information. However, because we changed the format of choice options to include the variance and total information (see Fig. 4A), these potential problems were removed from Study 2.

This consideration leads us to a revised version of H3:

H3r. : The extent to which the maximin and the egalitarian dimensions are compared directly in the discussion relates to shifts from an egalitarian to a maximin strategy.

3. Study 2

3.1. Method

3.1.1. Participants

A total of 166 student volunteers at a Japanese university

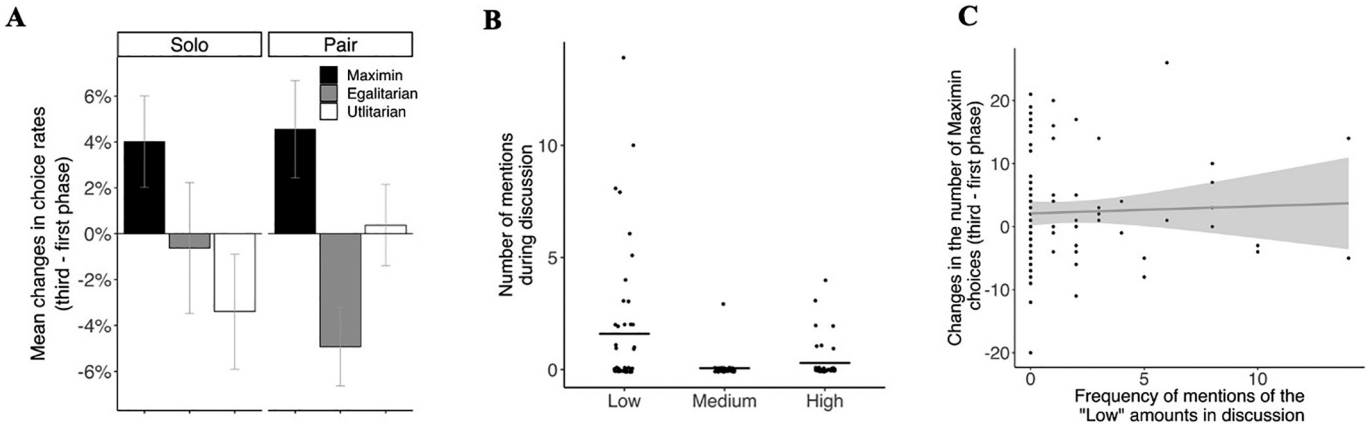


Fig. 3. Choice changes between the first and the third phases, and conversation patterns during the second phase of Study 1. (A) Mean changes in the choice rates of the three options between the first and the third phase by condition. Error bars represent SEM. (B) Number of times each amount (Low, Medium, and High) was mentioned by each pair. To take an example of the choice problem at the bottom half of Fig. 1A, mentions of 160, 170, and 290 (in yen) were counted as mentions of low amounts. Of the 50 pairs, three were removed from the conversation analysis because the voice recorder failed to function, leaving us with 47 pairs in total. Each point in the Low, Medium, and High amount columns corresponds to one pair (i.e., there are 47 dots each for the Low, Medium, and High amounts). The points were jittered to better show density. Crossbars represent mean number of mentions in the pair discussion. (C) Relationship between frequencies of mentions of "Low" amounts during pair discussion and incremental individual preference for the Maximin allocations. Shaded areas indicate a 95% confidence interval for fitted lines. Note that, for each frequency at the pair level (X-axis), there are two opinion-change scores by two members in the pair (Y-axis). The analysis reported in the main text took this nested structure into account.

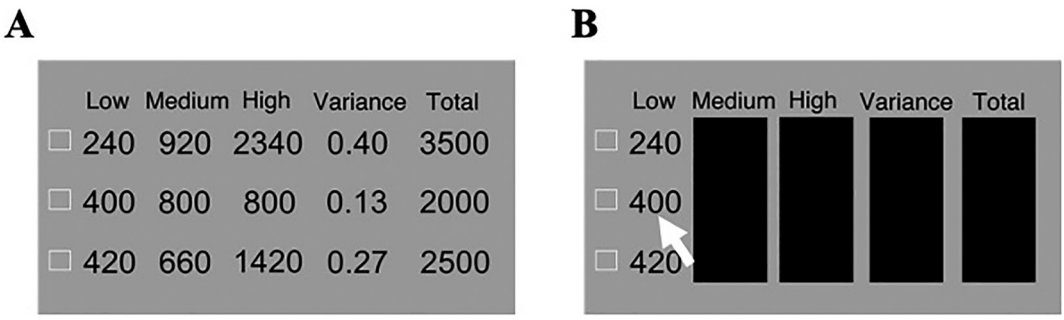


Fig. 4. Stimulus presentation in Study 2. (A) An example display of three choice options (in Yen) in Study 2. The choice problems were presented along with summary information ("Variance" and "Total"), in addition to the Low, Medium, and High amounts. Participants were instructed that numbers under the "Variance" label represented Gini coefficients used in economics that could range from 0 (perfect equality) to 1 (perfect inequality). (B) The Mouselab interface introduced in the second phase of Study 2. Numeric information was hidden behind boxes and became visible only when a participant moved the mouse pointer over a box and clicked.

participated in the experiment (91 male; M_{age} = 19.31, *SD* = 1.25). Minimum sample size was determined prior to the start of data collection as follows. In conversation analysis using a repeated measures ANOVA with five within factors (Low, Medium, High, Variance, and Total) given $\alpha = 0.05$ and power = 0.80, a sample size of 50 pairs allows us to detect an effect as small as 0.16 (a small-to-medium effect using Cohen's *f*). This led us to have at least 50 pairs in the Pair condition, and accordingly at least 50 individuals in the Solo condition. This yielded a total of at least 150 participants. No data analysis was conducted before we finished data collection. Informed consent was obtained from each participant using a consent form approved by the Institutional Review Board of the Center for Experimental Research in Social Sciences at the University.

3.1.2. Materials

As an internal replication of Study 1, we used the same 40 choice sets. In contrast with Study 1, however, the choice sets in Study 2 were presented along with summary information for each choice: Gini coefficient of distribution ("Variance") and the total amount of money to be allocated ("Total"), in addition to the original information (the low, medium, and high amounts: Fig. 4A). By including the summary information corresponding to the egalitarian dimension in the stimulus presentation, we could test whether the extent of direct comparisons recruited between the values for "Low" (which can be used to determine the maximin option) and "Variance" (which can be used to determine the egalitarian option) during discussion contributes to the increase in endorsements for the Maximin allocations after pair discussion (H3r).

In addition to the changes in stimulus presentation, we made two improvements to the second phase of Study 2. First, we introduced a mouse-tracing technique called Mouselab (Payne, Bettman, & Johnson, 1993) to the second phase, in which each option's numerical information was hidden behind a black box (Fig. 4B), and was only made visible when the participant hovered the mouse pointer over and clicked the box, and was hidden again when the participant moved the pointer out of the box. This attention-monitoring technique enabled us to compare individuals' information search patterns between the Pair condition and the Solo condition. Participants were free to view any information in any order. Second, for the results of Study 1, we speculated that using the fixed set of five problems for the second phase could have biased participants' choices in the third phase, as indicated by the seemingly peculiar decrease in the number of Utilitarian choices in the Solo condition (Fig. 3A). To control for the possibility of such potential confounds, in the second phase of Study 2, we randomly selected 12 problems out of the 40 problems for each pair or each participant (instead of using the pre-fixed five sets for all participants). In Study 2, we used a Microsoft Surface Pro 4 to present stimuli with a PsychoPy script (Peirce et al., 2019).

In the first and third phases, the column orders of Low, Medium and High amounts (L-M-H or H-M-L) were counterbalanced every 10 problems, and the column orders of Variance and Total were counterbalanced every 20 problems. In the second phase, the column orders of Low, Medium and High amounts were counterbalanced every 3 problems, and the column orders of Variance and Total were counterbalanced every 6 problems. Regardless of the column orders of Low, Medium, and High, participants were informed that person A (X, H) would receive the lowest outcome, B (Y, I) the middle outcome, and C (Z, J), the highest outcome.

3.1.3. Procedure

Most of the experimental procedure was the same as in Study 1. Participants made 40 allocation choices individually in the first and third phase. In the second phase, participants in the Pair condition (N = 102: 18 male-male pairs, 16 female-female pairs, and 17 mixed-gender pairs) made a consensus decision for each of the 12 problems, while those in the Solo condition (N = 64, 39 males) made decisions alone. Participants had to choose within 30 s for each of the 40 problems in the first and third phase. In the second phase, the time limit was longer, with a total of 20 min for the 12 problems.

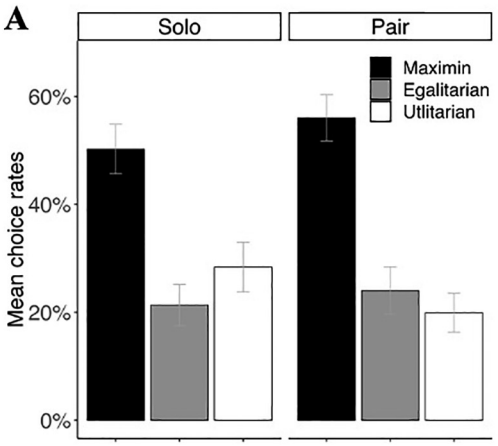
At the end of the experiment, participants answered a series of questions about risk preferences (Eckel & Grossman, 2002), and then received 1000 yen for participation and were dismissed. The recipients were also paid later according to the aforementioned procedure.

3.2. Results

3.2.1. Behavioral data

As in Study 1, we first analyzed pair decisions in the second phase. Fig. 5A displays decision rates for the Maximin, Egalitarian, or Utilitarian options. A multinomial logistic regression analysis (with the Maximin options as a reference and with varying intercepts for participants or pairs) again revealed that the Maximin options were chosen more frequently than the Egalitarian options ($\beta = -2.15$, 95% CI [−3.22, −1.16]) or the Utilitarian options ($\beta = -1.72$, 95% CI [−2.79, −0.70]). The difference between the Solo and the Pair conditions was not significant ($\beta = 0.11$, 95% CI [−1.28, 1.50] for the Egalitarian options; $\beta = -0.72$, 95% CI [−2.28, 0.71] for the Utilitarian options). Likewise, as seen in Fig. 5B, the social decision scheme matrix showed that the Egalitarian-Utilitarian pairs (bottom row) chose the Maximin option more frequently (0.44) compared to the two options of the pair members' original preferences (0.24, 0.33 respectively), which was not statistically predictable from the model that discussion would yield only either Egalitarian or Utilitarian consensus with equal probability (Bayes factor > 100; multinomial test).

Compared to the Solo condition, in the Pair condition the number of Egalitarian choices decreased between the first and third phases, while the number of Maximin choices increased (Fig. 6A; $\beta_{Pair*third\ phase} = 0.68$, 95% CI [0.45, 0.92] by a multinomial logistic regression analysis with the Egalitarian options as a reference and with varying intercepts



B

Pair decisions

Individual configurations	Maximin	Egalitarian	Utilitarian	Freq.
(Maximin, Maximin)	0.78	0.12	0.09	170
(Egalitarian, Egalitarian)	0.2	0.8	0	50
(Utilitarian, Utilitarian)	0.22	0.03	0.75	36
(Maximin, Utilitarian)	0.56	0.04	0.4	141
(Maximin, Egalitarian)	0.56	0.41	0.03	160
(Egalitarian, Utilitarian)	0.44	0.24	0.33	55

Fig. 5. Choice results in the second (deliberation) phase of Study 2. (A) Mean choice rates of the three options in the Solo and Pair conditions. Error bars represent SEM. (B) Social decision scheme matrix indicating how each pair aggregated initial individual preferences (first phase) into pair decisions (second phase) on the twelve problems that were repeated across both phases. See also Fig. S3 for an analysis using individual configurations based on participant's "behavioral type" (Egalitarian, Maximin, or Utilitarian type), instead of the choice-based configuration for each problem.

for participants; see Fig. S4 for the within-person choice stability of participants during the first phase). We also confirmed that this shift in behavioral preferences was observed not only in the 12 problems presented in the second phase but also in the remaining 28 problems not considered in the second phase ($\beta_{Pair*third\ phase}$ = 0.78, 95% CI [0.50, 1.06]) and was significant when gender was included as a covariate ($\beta_{Pair*third\ phase}$ = 0.68, 95% CI [0.45, 0.91]). These results replicated the results of Study 1, providing clear support for H1, H2, and H4.

3.2.2. Conversation data

The conversation pattern in the Pair condition was also replicated. The average number of words (nouns, verbs, adjectives, and adverbs) uttered during pair discussion in each group was 356.98 words. Pairs mentioned Low information most frequently, compared to any other type of information (β_{medium} = −0.79, 95% CI [−0.97, −0.63]; β_{high} = −0.70, 95% CI [−0.86, −0.55]; $\beta_{variance}$ = −0.63, 95% CI [−0.80, −0.47]; β_{total} = −0.73, 95% CI [−0.90, −0.57] by Poisson regression with varying intercepts for pairs).

We next examined how pair discussion promoted differentiation between the Maximin and the Egalitarian logics. We hypothesized that direct comparisons between the "Low" amount (the Maximin dimension) and the "Variance" information (the Egalitarian dimension) during discussion would promote shifts from an egalitarian to a maximin strategy (H3r). Thus, we used the frequency of consecutive mentions of "Low" and "Variance" in each dialogue as a measure of direct comparisons. Table 1 displays an example of dialogue. Here, "Low" and "Variance" were mentioned consecutively twice by the pair (one was "Low" followed by "Variance" and the other was "Variance" followed by "Low"). Fig. 6B displays relations between frequencies of such consecutive mentions of the "Low" and "Variance" information during pair discussion and incremental individual preference for the Maximin allocations. As seen in the graph, the greater the consecutive mentions during pair discussion, the greater the individual preference for the Maximin allocations after discussion. A linear regression with varying intercepts for pairs showed that the frequency of consecutive mentions of "Low" and "Variance" contributed to the increase in the number of

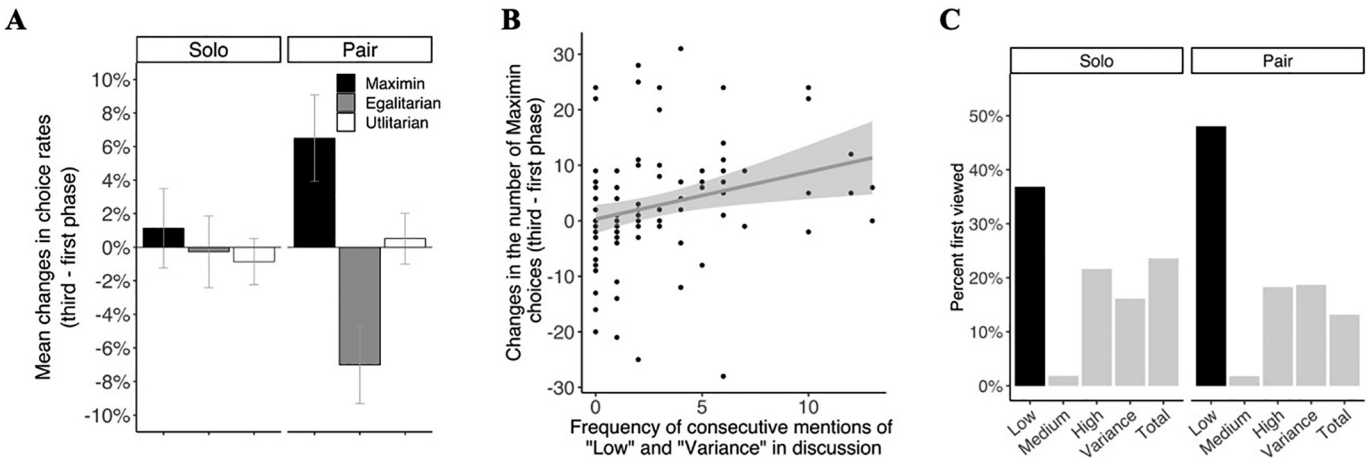


Fig. 6. Choice changes between the first and the third phases, conversation, and information search patterns during the second phase of Study 2. (A) Mean changes in the choice rates of the three options between the first and the third phase by condition. Error bars represent SEM. (B) Relationship between frequencies of consecutive mentions of the "Low" and "Variance" information during pair discussion (see the note accompanying Table 1 for definition) and incremental individual preference for Maximin allocations. Shaded areas indicate a 95% confidence interval for fit lines. Note that, for each frequency at the pair level (X-axis), there are two preference-change scores by two members in the pair (Y-axis). The analysis reported in the main text took this nested structure into account. (C) Percentages of first views (i.e., the first box which was clicked in the Mouselab interface during each of the 12 trials) collapsed across the second phase.

Table 1
An example of pair dialogue from Study 2.

No.	Person	content
1	A	How did you choose options in the first phase?
2	B	I made choices based on the **Low**. I prioritized to improve the welfare of the person who would receive the **Low** amount.
3	A	OK. My choices were based on the **Variance**, and I basically minimized the **Variance**. However, I reconsidered this rule when the **Low** was below 200 yen because I felt sorry for the poorest recipient.
4	B	I see. (dialogue continued)

Note. In this example case, four consecutive mentions are found in the order of occurrence: (a) "Low" and "Low" (LL), (b) "Low" and "Variance" (LV), (c) "Variance" and "Variance" (VV), and (d) "Variance" and "Low" (VL). The frequency of consecutive mentions of "Low" and "Variance" (LV + VL, the key explanatory variable in the regression analysis for Fig. 6B) was thus two, (b) + (d), and that for "Low" and "Low" and for "Variance" and "Variance" (see Table S2) was one, (a) and (c) respectively. We counted the number of consecutive mentions with the three possible patterns (LL, LV + VL, VV) in all dialogues bearing on 12 problems for each pair.

Maximin choices after pair discussion ($\beta_{Low-Variance}$ = 2.76, 95% CI [0.66, 4.85]). On the other hand, the frequency of repeated mentions of "Lows" ("Low" followed by "Low") and "Variances" ("Variance" followed by "Variance") did not have a meaningful relationship with the increase in Maximin choices (see Table S2). Thus, our revised hypothesis about the conversation effect (H3r) was supported.

3.2.3. Process-tracing data

Finally, using the Mouselab technique (Fig. 4B), we compared individuals' information search patterns between the Pair condition and the Solo condition. We conjectured that, compared to the Solo condition, participants in the Pair condition would focus on the minimum (Low) amounts more, anticipating that the concern for the poorest recipient would be a key topic during pair discussion. As shown in Fig. 6C, participants in the Pair condition, compared to participants in the Solo condition, were more likely to start information search during each of the 12 trials in the second phase by first clicking the Low box (β_{Pair} = 1.12, 95% CI [0.02, 2.25] by a logistic regression analysis with varying intercepts for participants and pairs; the display order of the amounts [L-M-H or H-M-L] was also modeled as a dummy variable). This pattern suggests that participants who engaged in paired discussion, compared with participants who faced the same decisions on their own, focused on the welfare of the worst-off recipient more selectively.

3.3. Discussion

Study 2 replicated results of Study 1, providing clear support for H1, H2, and H4. Study 2 also showed that the frequency of direct comparisons between the "Low" and the "Variance" information predicted the increase in Maximin preferences after pair discussion. This result supported H3r, suggesting that the comparison process during pair discussion is key for people to understand the maximin logic clearly.

In a follow-up to Study 2, we test the durability of these effects. If the cognitive changes in Study 2 proved long lasting, this would demonstrate that some egalitarian participants had genuinely internalized the maximin logic, which arguably had underlaid (and mistakenly manifested as) their initial egalitarian preferences. The follow-up study tested this internalization hypothesis using an online survey, in which participants from Study 2 were confronted again with the same decisions five months after Study 2.

4. Follow-up to Study 2

4.1. Method

4.1.1. Participants

Five months after Study 2, we contacted all 166 participants in Study 2 via e-mail to ask for their participation in an online study. Eighty-four of these participants (50 male; M_{age} = 19.86, SD = 1.28) agreed to participate in the online study. Informed consent was obtained from each participant using an online consent form approved by the Institutional Review Board of the Center for Experimental Research in Social Sciences at the university.

4.1.2. Materials

We used the same 40 choice sets as in Study 2. All the participants made the same 40 allocation choices to three unknown others (labeled persons E, F, and G) individually.

4.1.3. Procedure

Participants entered decisions in an online form using a smartphone or a personal computer. In case some participants had poor internet connections, we did not set time limits. We presented the choice problems just as in Study 2. After making choices for 40 allocation problems, the participants received a 1000-yen Amazon gift certificate for their participation. The recipients were also paid later as in Studies 1 and 2.

4.2. Results

Fig. 7 displays the mean choices of the 84 participants who responded to the online survey. Compared to the choices in the first phase of Study 2, participants who had been in the Pair condition retained their preference for the Maximin allocations, while those who had been in the Solo condition increased their preference for Egalitarian allocations. A multinomial logistic regression analysis with the fixed effect of phase (the first phase of Study 2 vs. follow-up) confirmed that participants in the Pair condition, compared to the participants in the Solo condition, retained their preference for the Maximin options over the Egalitarian options in the follow-up study ($\beta_{Pair*Study\ 3}$ = 1.10, 95% CI

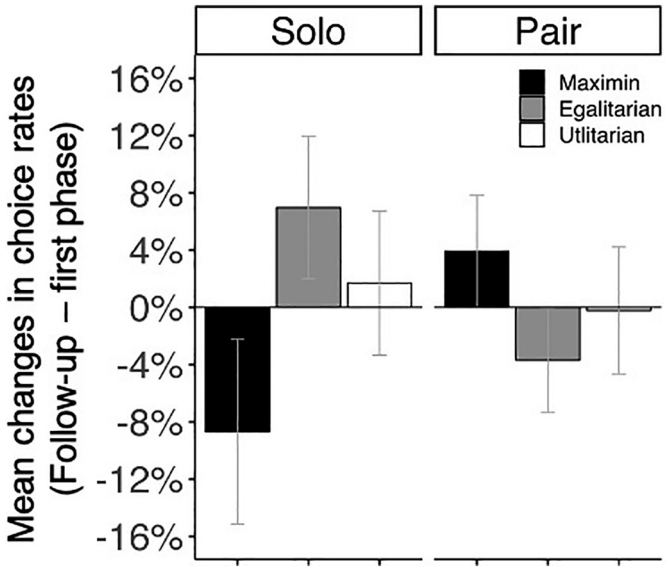


Fig. 7. Results from the follow-up to Study 2. The graph shows mean changes in the choice rates of the three options between the first phase of Study 2 and the follow-up (i.e., the follow-up minus the first phase of Study 2). Error bars represent SEM.

[0.78, 1.41] with varying intercepts for participants). We further confirmed that the simple effect was also credible in the Pair condition ($\beta_{Study\ 3}$ = 0.46, 95% CI [0.23, 0.69] by a multinomial logistic regression analysis with varying intercepts for participants).

We confirmed that there were no systematic differences in terms of age, gender, or risk preferences between those who participated in the follow-up and those who did not (see Table S3). Nevertheless, to avoid possible biased inferences caused by ignoring missing data (82 out of the 166 [49.4%] participants did not respond to the online study), we also imputed these missing data using a multiple imputation approach implemented in the R package mice v. 3.4.0 (van Buuren & Groothuis-Oudshoorn, 2011). As initially planned, we used participants' gender and risk preferences (Eckel & Grossman, 2002) to impute these missing data (see also Fig. S5 showing the relationship between risk preferences and distributive choices, which replicated the results of Kameda et al., 2016). The statistical results were unchanged by this multiple imputation approach ($\beta_{Pair*Study\ 3}$ = 0.78, 95% CI [0.57, 1.00] with varying intercepts for participants).

These results show that the preference shift towards the Maximin allocations and away from the Egalitarian allocations observed in Study 2 lasted at least five months after the experiment.

5. General discussion

Across two studies, participants who had discussed resource allocations in pairs (social deliberation) made more maximin choices and fewer egalitarian choices than participants who had reflected on the same problems on their own (solitary deliberation). A third study revealed that this shift was long-lasting, persisting five months after the experiment.

Conversation analyses revealed that the discussions focused on the outcomes of the poorest recipients, but that this focus wasn't sufficient to shift people from egalitarian to maximin preferences. Instead, it was the comparison between the main features of the egalitarian option (i.e. low variance) and the maximin option (i.e. relatively high allocation for the least well-off) that prompted this shift.

After dialogue, the participants' preferences shifted towards maximin even for choices that had not been part of the discussion, suggesting that they had paid attention to the logic underlying the maximin preference, and extrapolated it from the choices discussed to other choices as transferable knowledge (Gick & Holyoak, 1983). The fact that these changes persisted for at least five months also argues against lower level explanations. Interestingly, Huang, Kendrick, and Yu (2014) reported that preference changes induced by just observing the ratings of other participants, without being able to discuss with them, lasted only for a very short period of time (no more than three days), suggesting the importance of dialogue for creating a durable opinion shift (see also Broockman & Kalla, 2016).

In studies of group decision making, the effects of discussion are often evaluated by comparing pre- and post-discussion performance against objective benchmarks (e.g., known facts or mathematical demonstrations: Jayles et al., 2017; Laughlin, 2011; Trouche et al., 2014). By contrast, studies lacking such objective benchmarks often had recourse to properties of aggregated opinions, for example pointing out that discussion led to more homogeneous opinions (Himmelroos & Christensen, 2014), making it less likely that group decisions succumb to social-choice paradoxes like majority cycling (List, Luskin, Fishkin, & McLean, 2013).

In the present studies, we observed such homogenization, since the pre-existing maximin preference was strengthened by the discussion. However, if our account is correct, this shift also reflects an internal change for some participants, as they developed more coherent maximin preferences. We suggested that some of the preference for the egalitarian outcome actually reflected concerns for the less well-off, which led to egalitarian choices only because such choices often also increase the welfare of the less well-off. If this interpretation is correct, some participants had chosen egalitarian outcomes heuristically, but then realized during the discussion that these egalitarian choices would negatively affect the poorest. As a result, they developed more internally coherent preferences by adopting a maximin strategy.

We believe that controlled laboratory experiments such as the present ones are critical to systematically and rigorously examine people's distributive preference in the social context. However, we recruited only Japanese university students as participants, whose distributive preferences are far from reflecting the full range of preferences that could be observed in different social and cultural environments (e.g., Yarkoni, 2019). We also realize that more extensive analyses of conversation data with larger sample size would allow deeper insights into the power of dialogue to shape people's distributive preferences.

Finally, while our experiments have shown that many participants who appeared to hold egalitarian principles could be shifted by social deliberation towards maximin principles, some participants remained staunchly egalitarian in their preferences. It is possible that different cognitive mechanisms underpin these strict egalitarian preferences, making them less amenable to change through discussion. Future research addressing these mechanisms will provide further insights about why disputes about just distribution can sometimes be bitter, and simultaneously illuminate possible routes to overcome the social and political divides that are deepening around economic inequality.

6. Conclusion

In this paper, we have addressed whether deliberation—one of the most critical elements in democracy (Cohen, 1986; Elster, 1998; Habermas, 1997)—facilitates differentiation between maximin logic and egalitarian logic in people's allocation choices. Past attempts at improving the internal coherence of participants' political opinions through deliberation have only yielded small effects (Gastil, Black, & Moscovitz, 2008; Gastil & Dillard, 1999). By contrast, our effects concerning maximin preferences are much more robust. Our results suggest that people can form more “enlightened preferences”—that is, at the limit, “the preferences that people would have if their information were perfect” (Mansbridge, 1983, p.25)—not only by receiving more information (see, e.g., Bartels, 1990; Carpini & Keeter, 1996), but also by discussing with each other and paying more attention to the internal coherence of their own judgments, as observed in reasoning tasks (see, Mercier & Sperber, 2017).

Open practices

This study earned the Open Data badge for transparent practices. The data has been posted at https://osf.io/6f4jw/

Declaration of Competing Interest

This research was supported by the Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research JP16H06324, and Japan Science and Technology Agency CREST Grant JPMJCR17A4 (17941861) to Tatsuya Kameda, and the Japan Society for the Promotion of Science Grant-in-Aid for JSPS fellows JP18J21498 to Atsushi Ueshima. Support from CiSHub at the University of Tokyo is also appreciated. Hugo Mercier's work is supported by the ANR grant EUR FrontCog ANR-17-EURE-0017*.

The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2020.104067.

References

Bago, B., & De Neys, W. (2019). The smart system 1: Evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking and Reasoning, 25*, 257–299. https://doi.org/10.1080/13546783.2018.1507949.

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting minds. *Science, 329*, 1081–1085. https://doi.org/10.1126/science.1185718.

Bartels, L. (1990). *Public opinion and political interests.* Chicago: Paper presented at the Annual meeting of the Midwest Political Science Association April.

Baumard, N., André, J. B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences, 36*, 59–78. https://doi.org/10.1017/S0140525X11002202.

Broockman, D., & Kalla, J. (2016). Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science, 352*, 220–224. https://doi.org/10.1126/science.aad9713.

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*, 1–28. https://doi.org/10.18637/jss.v080.i01.

Burstein, P. (2003). The impact of public opinion on public policy: A review and an agenda. *Political Research Quarterly, 56*, 29–40. https://doi.org/10.1177/106591290305600103.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*, 1–67. https://doi.org/10.18637/jss.v045.i03.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software, 76*, 1–31. https://doi.org/10.18637/jss.v076.i01.

Carpini, M. X. D., & Keeter, S. (1996). *What Americans know about politics and why it matters.* Yale University Press.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics, 117*, 817–869. https://doi.org/10.1162/003355302760193904.

Claidière, N., Trouche, E., & Mercier, H. (2017). Argumentation and the diffusion of counter-intuitive beliefs. *Journal of Experimental Psychology: General, 146*, 1052–1066. https://doi.org/10.1037/xge0000323.

Cohen, J. (1986). An epistemic conception of democracy. *Ethics, 97*, 26–38. https://doi.org/10.1086/292815.

Core Team, R. (2017). *R: A language and environment for statistical computing.* (R Foundation for Statistical Computing).

Davis, J. H. (1973). Group decision and social interaction: A theory of social decision schemes. *Psychological Review, 80*, 97–125. https://doi.org/10.1037/h0033951.

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature, 446*, 794–796. https://doi.org/10.1038/nature05651.

Descartes, R. (1637/2006). *A Discourse on the method (I. Maclean, Trans.).* Oxford University Press.

Eckel, C. C., & Grossman, P. J. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior, 23*, 281–295. https://doi.org/10.1016/S1090-5138(02)00097-1.

Edwards, B. J., Williams, J. J., Gentner, D., & Lombrozo, T. (2019). Explanation recruits comparison in a category-learning task. *Cognition, 185*, 21–38. https://doi.org/10.1016/j.cognition.2018.12.011.

Elster, J. (1998). The market and the forum: Three varieties of political theory. In R. E. Goodin, & P. Pettit (Eds.). *Contemporary political philosophy: An anthology (2nd ed) (pp. 144–158).* Blackwell.

Engelmann, D., & Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American Economic Review, 94*, 857–869. https://doi.org/10.1257/0002828042002741.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics, 114*, 817–868. https://doi.org/10.1162/003355399556151.

Fehr, E., & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism –experimental evidence and new theories. In S.-C. Kolm, & J. M. Ythier (Eds.). *Handbook of the economics of giving, altruism and reciprocity (pp. 615–691).* Elsevier.

Frohlich, N., & Oppenheimer, J. A. (1992). *Choosing justice: An experimental approach to ethical theory.* University of California Press.

Frohlich, N., Oppenheimer, J. A., & Eavey, C. L. (1987). Laboratory results on Rawls's distributive justice. *British Journal of Political Science, 17*, 1–21. https://doi.org/10.1017/S0007123400004580.

Gastil, J., Black, L., & Moscovitz, K. (2008). Ideology, attitude change, and deliberation in small face-to-face groups. *Political Communication, 25*, 23–46. https://doi.org/10.1080/10584600701807836.

Gastil, J., & Dillard, J. P. (1999). Increasing political sophistication through public deliberation. *Political Communication, 16*, 3–23. https://doi.org/10.1080/105846099198749.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170. https://doi.org/10.1016/S0364-0213(83)80009-3.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1–38. https://doi.org/10.1016/0010-0285(83)90002-6.

Habermas, J. (1997). Popular sovereignty as procedure. In J. Bohman, & W. Regh (Eds.). *Deliberative democracy: Essays on reason and politics (pp. 35–65).* MIT Press.

Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review, 112*, 494–508. https://doi.org/10.1037/0033-295X.112.2.494.

Himmelroos, S., & Christensen, H. S. (2014). Deliberation and opinion change: Evidence from a deliberative mini-public in Finland. *Scandinavian Political Studies, 37*, 41–60. https://doi.org/10.1111/1467-9477.12013.

Hsu, M., Anen, C., & Quartz, S. R. (2008). The right and the good: Distributive justice and neural encoding of equity and efficiency. *Science, 320*, 1092–1095. https://doi.org/10.1126/science.1153651.

Huang, Y., Kendrick, K. M., & Yu, R. (2014). Conformity to the opinions of other people lasts for no more than 3 days. *Psychological Science, 25*, 1388–1393. https://doi.org/10.1177/0956797614532104.

Ishida, M. (2018). *RMeCab: Interface to MeCab (v. 0.99999).*

JASP Team (2020). *JASP (v. 0.12.2).* https://jasp-stats.org/.

Jayles, B., Kim, H., Escobedo, R., Cezerad, S., Blanchet, A., Kameda, T., ... Theraulaz, G. (2017). How social information can improve estimation accuracy in human groups. *Proceedings of the National Academy of Sciences of the United States of America, 114*, 12620–12625. https://doi.org/10.1073/pnas.1703695114.

Jensen, K. B. (2016). Two-step and multistep flows of communication. In K. B. Jensen, E. W. Rothenbuhler, J. D. Pooley, & R. T. Craig (Eds.). *The international encyclopedia of communication theory and philosophy (pp. 1–11).* Wiley.

Kameda, T., Inukai, K., Higuchi, S., Ogawa, A., Kim, H., Matsuda, T., & Sakagami, M. (2016). Rawlsian maximin rule operates as a common cognitive anchor in distributive justice and risky decisions. *Proceedings of the National Academy of Sciences of the United States of America, 113*, 11817–11822. https://doi.org/10.1073/pnas.1602641113.

Kameda, T., Tindale, R., & Davis, J. (2003). Cognitions, preferences, and social sharedness: Past, present, and future directions in group decision making. In S. Schneider, & J. Shanteau (Eds.). *Emerging perspectives on judgment and decision research* (pp. 458–485). Cambridge University Press. https://doi.org/10.1017/CBO9780511609978.016.

Kameda, T., Tsukasaki, T., Hastie, R., & Berg, N. (2011). Democracy under uncertainty: The wisdom of crowds and the free-rider problem in group decision making. *Psychological Review, 118*, 76–96. https://doi.org/10.1037/a0020699.

Kerr, N. L., MacCoun, R. J., & Kramer, G. P. (1996). Bias in judgment: Comparing individuals and groups. *Psychological Review, 103*, 687–719. https://doi.org/10.1037/0033-295X.103.4.687.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 107–118. https://doi.org/10.1037/0278-7393.6.2.107.

Kudo, T. (2018). *MeCab: Yet another part-of-speech and morphological analyzer (v. 0.996).*

Kugler, T., Kausel, E. E., & Kocher, M. G. (2012). Are groups more rational than individuals? A review of interactive decision making in groups. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*, 471–482. https://doi.org/10.1002/wcs.1184.

Laughlin, P. R. (2011). *Group problem solving.* Princeton University Presshttps://doi.org/10.1515/9781400836673.

Laughlin, P. R., & Ellis, A. L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology, 22*, 177–189. https://doi.org/10.1016/0022-1031(86)90022-3.

List, C., Luskin, R. C., Fishkin, J. S., & McLean, I. (2013). Deliberation, single-peakedness, and the possibility of meaningful democracy: Evidence from deliberative polls. *The Journal of Politics, 75*, 80–95. https://doi.org/10.1017/S0022381612000886.

Mansbridge, J. J. (1983). *Beyond adversary democracy.* University of Chicago Press.

Masclet, D., Colombier, N., Denant-Boemont, L., & Loheac, Y. (2009). Group and individual risk preferences: A lottery-choice experiment with self-employed and salaried workers. *Journal of Economic Behavior and Organization, 70*, 470–484. https://doi.org/10.1016/j.jebo.2007.11.002.

Mercier, H. (2016a). Confirmation bias - myside bias. In R. F. Pohl (Ed.). *Cognitive illusions: intriguingphenomena in thinking, judgment and memory* (pp. 99–114). Psychology Press. https://doi.org/10.4324/9781315696935.

Mercier, H. (2016b). The argumentative theory: Predictions and empirical evidence. *Trends in Cognitive Sciences, 20*, 689–700. https://doi.org/10.1016/j.tics.2016.07.001.

Mercier, H., & Landemore, H. (2012). Reasoning is for arguing: Understanding the successes and failures of deliberation. *Political Psychology, 33*, 243–258. https://doi.org/10.1111/j.1467-9221.2012.00873.x.

Mercier, H., & Sperber, D. (2017). *The enigma of reason.* Harvard University Press.

Mitchell, G., Tetlock, P. E., Mellers, B. A., & Ordóñez, L. D. (1993). Judgments of social justice: Compromises between equality and efficiency. *Journal of Personality and Social Psychology, 65*, 629–639. https://doi.org/10.1037/0022-3514.65.4.629.

Moshman, D., & Geil, M. (1998). Collaborative reasoning: Evidence for collective rationality. *Thinking and Reasoning, 4*, 231–248. https://doi.org/10.1080/135467898394148.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*, 175–220. https://doi.org/10.1037/1089-2680.2.2.175.

Ogawa, A., Ueshima, A., Inukai, K., & Kameda, T. (2018). Deciding for others as a neutral party recruits risk-neutral perspective-taking: Model-based behavioral and fMRI experiments. *Scientific Reports, 8*, 12857. https://doi.org/10.1038/s41598-018-31308-6.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker.* Cambridge University Press.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*, 195–203. https://doi.org/10.3758/s13428-018-01193-y.

Rawls, J. (1971). *A theory of justice.* Harvard University Press.

Shaw, A. (2013). Beyond "to share or not to share". *Current Directions in Psychological Science, 22*, 413–417. https://doi.org/10.1177/0963721413484467.

Starmans, C., Sheskin, M., & Bloom, P. (2017). Why people prefer unequal societies. *Nature Human Behaviour, 1*, 1–7. https://doi.org/10.1038/s41562-017-0082.

Stone, E. R., Yates, A. J., & Caruthers, A. S. (2002). Risk taking in decision making for others versus the self. *Journal of Applied Social Psychology, 32*, 1797–1824. https://

doi.org/10.1111/j.1559-1816.2002.tb00260.x.
Sunstein, C. R. (2002). The law of group polarization. *Journal of Political Philosophy, 10*, 175–195. https://doi.org/10.1111/1467-9760.00148.
Tesser, A. (1978). Self-generated attitude change. In L. Berkowitz (Ed.). *Advances in experimental social psychology (pp. 289–338)*Academic Presshttps://doi.org/10.1016/S0065-2601(08)60010-6.
Tindale, R. S., Smith, C. M., Thomas, L. S., Filkins, J., & Sheffey, S. (1996). Shared representations and asymmetric social influence processes in small groups. In E. H. Witte, & J. H. Davis (Eds.). *Understanding group behavior: Consensual action by small groups (pp. 81–103)*. Lawrence Erlbaum.
Trouche, E., Sander, E., & Mercier, H. (2014). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General, 143*, 1958–1971. https://doi.org/10.1037/a0037099.
Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* Springer.
Yarkoni, T. (2019). *The Generalizability Crisis.* https://doi.org/10.31234/osf.io/jqw35 November 22.