LEBS
Vol. 12 No. 2 (2021) 34–38
Letters on Evolutionary Behavioral Science

Inequality Biases Third-Party Evaluation of Decision-Making for Others

Kiri Kuroda[1,2,3,*], Yoshimatsu Saito[4]

[1]Department of Social Psychology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

[2]Japan Society for the Promotion of Science, 5-3-1 Kojimachi, Chiyoda-ku, Tokyo 102-0083, Japan

[3]Institute for Research in Business and Economics, Faculty of Economics, Meiji Gakuin University, 1-2-37 Shirokanedai, Minato-ku, Tokyo 108-8636, Japan

[4]BrainPad Inc., 3-2-10 Shirokanedai, Minato-ku, Tokyo 108-0071, Japan

[*]Author for correspondence (kuroda.kiri@gmail.com)

People often need to make risky decisions for others, especially in policymaking, where a single decision can affect the welfare of a number of people. Given that risky decisions can yield variable outcomes and that people often evaluate policies after knowing the outcomes, the same risky policy can be evaluated differently depending on its outcome. Nevertheless, very little is known about how people make third-party evaluations of risky policies. Because people are sensitive to inequality among others, we predicted that the same policy would be evaluated more negatively if it leads to inequality rather than other outcomes. To examine this, we conducted a scenario experiment on risky and sure policies and investigated whether people's distributive preferences moderated policy evaluation. We show that participants rated the risky policy lower when it yielded unequal situations between the recipients. Interestingly, participants did not evaluate the risky policy higher than the sure policy even when the risky policy yielded more desirable outcomes. In addition, participants who preferred sure distributions as decision makers or recipients showed the inequality aversion, whereas participants who preferred risky distributions showed no such pattern. Our results suggest that policy evaluation may be susceptible to the risks and inequality of outcomes among recipients.

Keywords
decision-making for others, inequality aversion, risk, outcome bias, distribution, policymaking

Introduction

People often have to make choices that affect the well-being of another person. Doctors may be required to choose between a less burdensome treatment with a higher risk of recurrence and a more burdensome treatment with a lower risk. Financial advisors need to recommend investment options to their clients while considering the balance between risk and return. As shown by these examples, decision-making for another person often involves uncertainty about outcomes. Concerning such risky decisions for another person, Hsee and Weber (1997) showed that decision makers tend to assume that another person is more tolerant toward risks than themselves, and Ogawa et al. (2018) suggested that people make less risk-averse decisions for another person than for themselves.

Risky decisions can have great impacts on the well-being of large numbers of people. Such situations are often found in policymaking, where a single risky decision can affect the well-being of many citizens. For example, Shenhav and Greene (2010) claimed that policy decisions involve tradeoffs between the magnitude of benefits (e.g., the number of beneficiaries or rescued people) and the probability of the desired outcomes. In fact, in their fMRI experiment using moral-dilemma scenarios, decision-making for a number of others recruited neural circuits adapted for risky decision-making. This result suggests that policymakers view policymaking as a risky decision for others.

As described above, decision-making for others has been well researched from the perspective of decision makers (e.g., Jung et al., 2013; Mengarelli et al., 2014; Polman, 2012). However, very little is known about how third parties think about the justifiability of decisions, although such judgments are essential, especially at the policy level.

In investigating how people evaluate risky policies, attention should be paid to two points. Firstly, policies are often assessed after knowing their outcomes. If the policy involves risks and yields variable outcomes, the same policy could be evaluated differently depending on its outcome. In fact, in moral decision-making, people have been known to change their third-party evaluations based on the outcome of decisions. This cognitive tendency has been called outcome bias (e.g., Baron & Hershey, 1988). Outcome biases may also arise in the evaluation of risky policies.

Secondly, unlike in interpersonal relations, risky decisions for others at the policy level can cause another social problem: inequality among recipients. If a policy involves some risks and affects the recipients independently, the policy can yield different consequences for different people. Such inequality caused by a risky policy would bias the evaluation of the policy itself because humans are averse not only to unequal outcomes between self and others (inequality aversion: Fehr & Schmidt, 1999) but also to unequal situations between others as a third party (Hsu et al., 2008; Zhong et al., 2016). If so, in order to win the support of citizens, policymakers may need to make sure that their risky policies lead to as little inequality as possible while increasing the overall wealth of society.

doi: 10.5178/lebs.2021.87
Received 15 June 2021.
Accepted 16 June 2021.
Published online 10 July 2021.

Here we conducted a scenario experiment to investigate how people make third-party evaluations of a risky policy of monetary distribution after knowing its consequences. We also measured participants' distributive preferences as decision makers or recipients to exploratorily examine whether these preferences moderate the evaluation of the risky policy. If a moderation effect is found, it would provide additional support that equitable outcomes among recipients play an important role in policy evaluation.

Methods

Participants

Two hundred and fifty-one students at Hokkaido University and the University of Tokyo participated in the experiment (Hokkaido: 79 men and 51 women; $M_{age} \pm SD_{age}$, 19.7 ± 1.2 years; Tokyo: 78 men and 43 women; 22.0 ± 2.1 years). Using an R package (Champely, 2020), we found that the minimum sample size was 175 to detect a medium effect ($\eta^2 = .06$ or Cohen's $f = 0.253$) with 80% power using a 5% significance level in a one-way ANOVA, the main analysis in this study (see Results for details). This means that the sample size was sufficient for this experiment.

The experiment was approved by the Institutional Review Boards of the Center for Experimental Research in Social Sciences at Hokkaido University and the Graduate School of Humanities and Sociology at the University of Tokyo. All participants gave written informed consent before the experiment and were compensated for their participation with 1,000 JPY.

Experimental procedure

(a) Overview

Each participant was called to the laboratory and asked to read two scenarios about monetary distribution and evaluate them. For one scenario, participants rated distributive policies as a third party (see *Policy evaluation* below). For the other scenario, participants rated monetary distribution as decision makers and recipients (see *Distributive preferences* below). The presentation order of the two scenarios was counterbalanced across participants. The experiment was performed using Qualtrics, a platform for online surveys.

(b) Policy evaluation

Participants read a scenario about Person A deciding on a distribution policy for Persons B and C, whom Person A has never met. It was emphasized to participants that Person A only makes the decision as a policymaker and that distributions would be paid out of a public budget, not from Person A's pocket.

In this scenario (hereafter the policy-evaluation scenario), Person A has to choose between a sure policy and a risky but potentially more profitable policy for Persons B and C. The sure policy guarantees each person 500,000 JPY, whereas the risky policy lets each person draw from a lottery with a 50−50 chance of winning 200,000 or 1,800,000 JPY. Note that, under the risky policy, each person draws from the lottery independently. Thus, the outcomes of the lotteries can be different between the recipients, and this is key in our experiment.

After reading the policy-evaluation scenario, participants were asked to assume that Person A chose the risky policy. Participants were then randomly informed of one of the possible consequences of the policy according to the following four conditions (Fig. 1a): (i) Success condition ($n = 63$): Both persons received 1,800,000 JPY; (ii) Failure condition ($n = 63$): Both persons received 200,000 JPY; (iii) Inequality condition ($n = 61$): Person B received 1,800,000 JPY, and Person C received 200,000 JPY; (iv) Control condition ($n = 64$): No information about the outcome was given to participants.

Participants next rated Person A's choice of the risky policy on a 6-point scale (1: *Very bad*; 6: *Very good*). Participants were also asked to assume that Person A adopted the sure policy (Fig. 1a) and to evaluate the choice on the 6-point scale. The order of rating the two policies was counterbalanced across participants.

(c) Distributive preferences

Participants read another scenario about a distribution decision in which they or their own interests were involved. In this scenario (hereafter the experimental-reward scenario), participants were asked to imagine that they participate in an experiment in which a randomly selected participant decides the rewards for the others. The decision maker has to choose between a sure option and a risky but possibly gainful option. The sure option promises each participant 800 JPY, whereas the risky option lets each participant draw from a lottery with a 50-50 chance of winning 200 or 2,000 JPY. If the decision maker chooses the risky option, each participant draws from the lottery individually. Thus, the outcomes of the lotteries can vary among participants as in the policy-evaluation scenario.

After reading the experimental-reward scenario, participants were asked to assume that they were appointed as decision maker. Participants then rated how likely they would be to choose either option on a 4-point scale (1: *Definitely choose the sure option*; 4: *Definitely choose the risky option*). Participants also reported how much they would want the decision maker to choose either option on the same 4-point scale, assuming that they were put in the position of receivers. The order of these items was counterbalanced across participants.

Results

No significant differences in the policy evaluation between the universities

Participants in the experiment were recruited from Hokkaido University and the University of Tokyo, and the difference in universities might have affected the policy evaluation.

To check whether there were any differences in the policy evaluation between the universities, we analyzed participants' ratings for the risky policy with a 4 (conditions: between-participants) × 2 (universities: between-participants) ANOVA. This ANOVA yielded no significant effects including the universities [university: $F(1, 243) = 0.80$, $p = .373$, $\eta_p^2 = 3.27e{-}03$; Condition × University: $F(3, 243) = 0.70$, $p = .553$, $\eta_p^2 = 8.57e{-}03$; Fig. 1b, left side]. For the sure policy, no significant difference in the rating was observed between the universities [$M_{Hokkaido} = 4.22$, $M_{Tokyo} = 4.31$; Welch's $t(237.33) = -0.68$, $p = .499$, Cohen's $d = -0.09$; Fig. 1b, right side].

These results show that there were no university biases

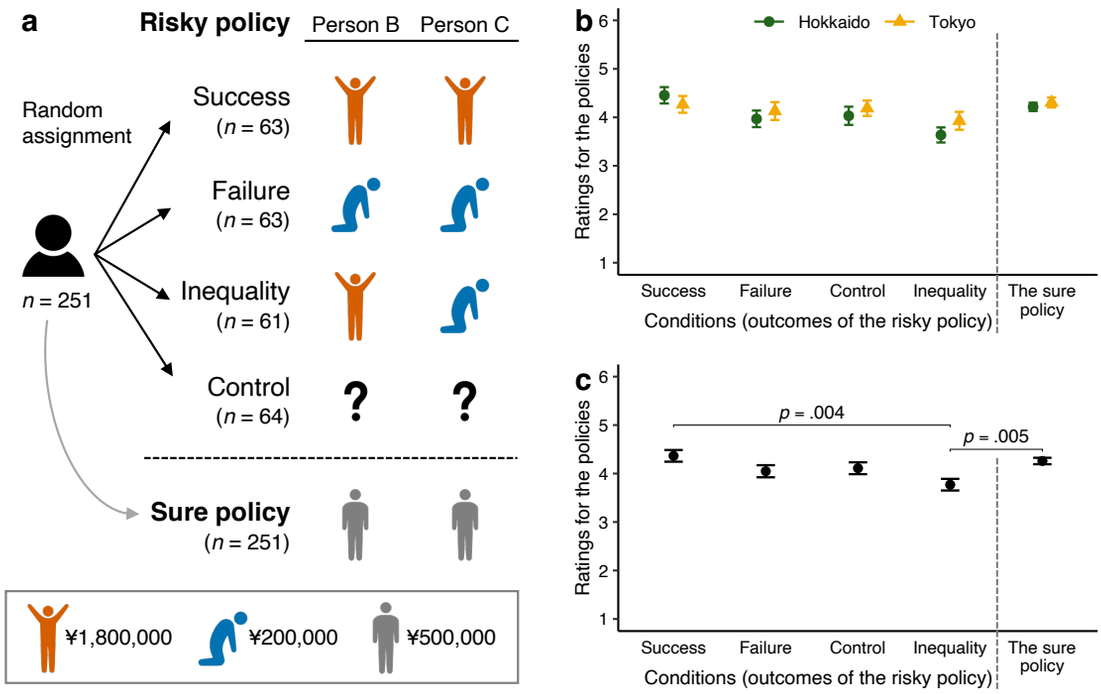

Figure 1. Experimental Conditions and Outcome Bias in the Policy Evaluation

Note. (a) Assignment to the conditions in the policy-evaluation scenario. Participants were randomly informed of one of the possible outcomes of the risky policy and then rated the policy. Participants also rated the sure policy regardless of the condition. (b) Ratings for the policies as a function of the conditions, policies, and universities (mean ± *SEM*). No significant differences were observed between the universities. (c) Ratings for the policies as a function of the conditions and policies (mean ± *SEM*). The risky policy was rated significantly lower in the inequality condition than in the success condition. The risky policy in the inequality condition was also rated significantly lower than the sure policy. In this figure, the ratings for the sure policy are collapsed across the four conditions.

in the policy evaluation. Therefore, we merged data from the two universities in the following analyses.

The unequal outcome biased the evaluation of the risky policy

To examine whether outcome biases occurred in the policy evaluation, we first analyzed participants' ratings for the risky policy with a one-way ANOVA with a between-participants factor of condition. The main effect of condition was significant, $F(3, 247) = 3.96$, $p = .009$, $\eta^2 = .05$. Pairwise comparisons also revealed a significant difference between the success and inequality conditions [$M_{success} = 4.37$, $M_{inequality} = 3.77$; $t(247) = -3.43$, Bonferroni corrected $p = .004$, $\eta_p^2 = .05$; Fig. 1c, left side]. No significant differences were observed for the other pairs of conditions, Bonferroni corrected $ps > .306$.

We next compared the policy evaluation between the risky and sure policies. Paired *t*-tests showed that, only in the inequality condition, the risky policy was rated significantly lower than the sure policy [$M_{risky} = 3.77$, $M_{sure} = 4.44$; $t(60) = 3.42$, Bonferroni corrected $p = .005$, Cohen's $d = 0.44$; Fig. 1c]. No significant difference was found between the policies in the success condition, [$M_{risky} = 4.37$, $M_{sure} = 4.22$; $t(62) = -0.65$, Bonferroni corrected $p = .999$, Cohen's $d = -0.08$] or in the other conditions [failure: $M_{risky} = 4.05$, $M_{sure} = 4.17$; $t(62) = 0.56$, Bonferroni corrected $p = .999$, Cohen's $d = 0.07$; control: $M_{risky} = 4.11$, $M_{sure} = 4.20$; $t(63) = 0.43$, Bonferroni corrected $p = .999$, Cohen's $d = 0.05$].

These results suggest that (i) the risky policy was not rated higher than the sure policy even if the risky policy was successful, and (ii) the risky policy was rated lower when the policy resulted in inequality.

Classifying participants as risky-type or sure-type by their distributive choices

To further investigate whether distributive preferences moderated the outcome bias in the policy evaluation, we first classified participants by their distributive choices in the experimental-reward scenario. In this classification, we defined ratings greater than the midpoint (i.e., 3 or 4) as choosing the risky option, and ratings less than the midpoint (i.e., 1 or 2) as choosing the sure option. Figure 2a shows that participants tended to choose the same option as a decision maker and as a receiver, $\chi^2(1) = 55.33$, $p = 1.018e{-}13$, Cramer's $V = .48$. Binomial tests also showed that more than half of participants preferred the sure option to the risky option in both positions, decision maker: 63.7% vs. 36.3%, $p = 1.577e{-}05$; receiver: 62.5% vs. 37.5%, $p = 8.409e{-}05$.

Since there was a strong association of preferences between the positions (i.e., decision-maker and receiver), we decided to categorize participants into three types (Fig. 2a): (i) the risky-type ($n = 62$: the orange cell), who preferred the risky option in both positions; (ii) the sure-type ($n = 128$: the blue cell), who preferred the sure option in both positions; (iii) the other-type ($n = 61$: the grey cells), who chose different options across the positions. Because we wanted to focus on differences between the risky- and sure-type preferences, the 61 other-type

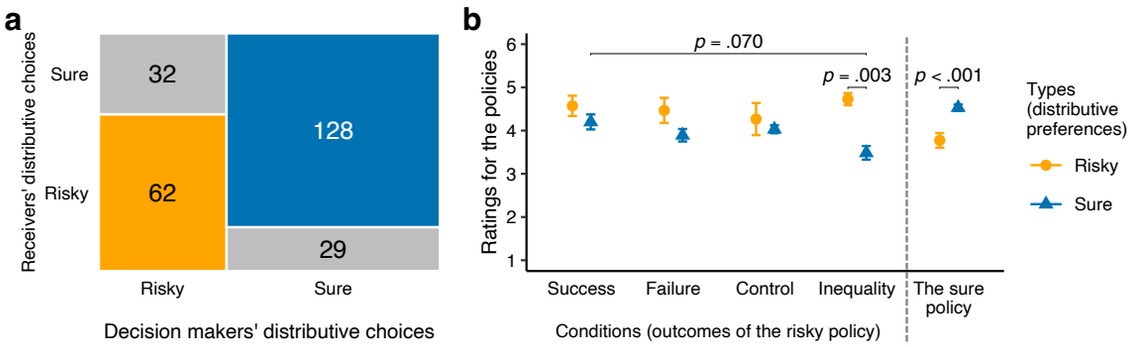

Figure 2. Moderation Effect of Participants' Distributive Preferences on the Policy Evaluation

Note. (a) Choice frequency of the risky and sure options in the experimental-reward scenario as a function of the participants' positions. The number in each cell and the size of each cell indicate the number of participants. If participants chose the same option across positions, participants were classified as preferring the option. (b) Ratings for the policies as a function of the conditions, policies, and distributive preferences (mean ± *SEM*). Only in the inequality condition, the sure-type participants rated the risky policy significantly lower than the risky-type participants. The sure-type participants tended to rate the risky policy lower in the inequality condition than in the success condition, whereas the risky-type participants showed no such differences in rating for the risky policy between the conditions. A significant difference was also observed in the rating for the sure policy between the two types of participants. In this figure, the ratings for the sure policy are collapsed across the four conditions.

participants were excluded from the analysis below.

Distributive preferences moderated the outcome bias in the policy evaluation

We then compared the evaluation of the policies between the risky- and sure-type participants. The sure policy was rated higher by the sure-type participants than the risky-type participants, $M_{\text{sure-type}} = 4.53$, $M_{\text{risky-type}} = 3.77$; Welch's $t(85.64) = -4.12$, $p = 8.738\text{e}{-05}$, Cohen's $d = -0.74$; Fig. 2b, right side. On the other hand, the risky policy was rated lower by the sure-type participants than the risky-type participants, $M_{\text{sure-type}} = 3.88$, $M_{\text{risky-type}} = 4.50$; Welch's $t(96.63) = 3.91$, $p = 1.718\text{e}{-04}$, Cohen's $d = 0.66$; Fig. 2b, left side.

We further analyzed the policy evaluation of the risky policy with a 4 (conditions: between-participants) × 2 (risky- and sure-types: between-participants) ANOVA. This ANOVA yielded significant main effects of condition and type, condition: $F(3, 182) = 2.97$, $p = .033$, $\eta_p^2 = .05$; type: $F(1, 182) = 15.59$, $p = 1.13\text{e}{-04}$, $\eta_p^2 = .08$. The ANOVA also revealed a marginally significant interaction between condition and type, $F(3, 182) = 2.17$, $p = .0998$, $\eta_p^2 = .03$.

Three interesting results were shown by pairwise comparisons (Fig. 2b, left side). Firstly, only in the inequality condition, the sure-type participants rated the risky policy lower than the risky-participants did, $M_{\text{sure-type}} = 3.48$, $M_{\text{risky-type}} = 4.73$; $t(182) = 3.90$, Bonferroni corrected $p = .003$, $\eta_p^2 = .08$. Secondly, the sure-type participants tended to rate the risky policy lower in the inequality condition than in the success condition, $M_{\text{inequality}} = 3.48$, $M_{\text{success}} = 4.20$; $t(182) = -2.94$, Bonferroni corrected $p = .070$, $\eta_p^2 = .05$. Thirdly, for the risky-type participants, no such difference was observed between the inequality and success conditions, $M_{\text{inequality}} = 4.73$, $M_{\text{success}} = 4.57$; $t(182) = 0.46$, Bonferroni corrected $p = .999$, $\eta_p^2 = 1.15\text{e}{-03}$. Taken together, the sure-type participants responded negatively to the unequal outcome caused by the risky policy, whereas the risky-type participants showed no such bias.

Discussion

Participants responded negatively as third parties to the unequal outcome caused by the risky policy (Fig. 1c), which is consistent with previous findings on third-party inequality aversion (Hsu et al., 2008; Zhong et al., 2016). In addition, this response was moderated by participants' distributive preferences as decision makers and recipients (Fig. 2b). These results show that participants emphasized equality rather than the expected total amount of wealth in the policy evaluation.

Interestingly, participants did not evaluate the risky but profitable policy higher than the sure policy even when the risky policy yielded more desirable outcomes (i.e., the success condition: Fig. 1c). This result would seem to reflect humans' risk-reduction notion about social distribution, and previous anthropological research has suggested that this risk-reduction notion has been shaped evolutionarily. For example, in a hunter-gatherer society, hunted game (e.g., peccary, monkey, deer) tends to be shared more widely beyond kinship than collected resources do (e.g., vegetables, fruits) (Kaplan et al., 1985, 1990). Kaplan et al. (1985) argued that this difference in resource sharing arise from uncertainty in resource acquisition and that social sharing functions as a collective risk-reduction device. In fact, such a collective risk-hedge is still found as insurance systems in modern society. If coping with uncertainty has been a recurrent adaptive task for humans, people may expect economic policies to act as a social safety net, rather than to increase the overall wealth of society.

However, this study has some limitations. Firstly, our experiment created a situation where participants evaluate policies as a pure third party, but in a real situation, people often assess policies while simultaneously being beneficiaries. Future studies should examine how beneficiaries evaluate a risky policy, especially when they benefit from the policy more or less than others. Secondly, in our experimental paradigm, we cannot strictly distinguish whether people were concerned about disparity between the two recipients (inequality aversion: e.g., Hsu

et al., 2008; Zhong et al., 2016) or the worst position itself, i.e., Person C in the policy-evaluation scenario (maximin concern: e.g., Charness & Rabin, 2002; Kameda et al., 2016). Finally, the policy in our experiment could be seen as too risky or unrealistic, and participants could have differently responded to more balanced policies. To address these points, future studies need to refine the experimental design by parametrically varying the magnitude of return and risk of policies.

Risky decisions for others have impacts on a number of recipients at the policy level. Unlike economic decision-making for oneself or another person (e.g., investment or gambling), such policy decisions can cause inequality among beneficiaries. In addition, policy decisions for resource allocation have often been expected to function as a collective hedge against risks rather than to increase the total wealth of society. Thus, from an evolutionary perspective, it may not be necessarily appropriate to frame policymaking (i.e., decision-making for a number of others) as simple risky-decisions to maximize the total wealth (cf. Shenhav & Greene, 2010). Our results would seem to suggest that policymakers must take into account that people are sensitive to risks and disparities as a third party in order to increase the legitimacy of their policies.

Acknowledgment

This study was supported by Japan Society for the Promotion of Science (JP21J30006). We thank Yohsuke Matsumoto, Shunya Sato, and Izumi Hirayama for their support in data collection. Support from the Center for Experimental Research in Social Sciences at Hokkaido University and the UTokyo Center for Integrative Science of Human Behavior is also appreciated.

Supplementary Material

The data and code that support the findings of this study have been deposited in Open Science Framework (https://osf.io/2d7rh/).

References

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, *54*(4), 569–579. https://doi.org/10.1037/0022-3514.54.4.569

Champely, S. (2020). *pwr: Basic functions for power analysis*. (R package version 1.3-0) [Computer software]. https://CRAN.R-project.org/package=pwr

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, *117*(3), 817–869. https://doi.org/10.1162/003355302760193904

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868. https://doi.org/10.1162/003355399556151

Hsee, C. K., & Weber, E. U. (1997). A fundamental prediction error: Self–others discrepancies in risk preference. *Journal of Experimental Psychology: General, 126*(1), 45–53. https://doi.org/10.1037/0096-3445.126.1.45

Hsu, M., Anen, C., & Quartz, S. R. (2008). The right and the good: Distributive justice and neural encoding of equity and efficiency. *Science*, *320*(5879), 1092–1095. https://doi.org/10.1126/science.1153651

Jung, D., Sul, S., & Kim, H. (2013). Dissociable neural processes underlying risky decisions for self versus other. *Frontiers in Neuroscience*, *7*, 15. https://doi.org/10.3389/fnins.2013.00015

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*(42), 11817–11822. https://doi.org/10.1073/pnas.1602641113

Kaplan, H., Hill, K., Cadeliña, R. V., Hayden, B., Hyndman, D. C., Preston, R. J., Smith, E. A., Stuart, D. E., & Yesner, D. R. (1985). Food sharing among Ache foragers: Tests of explanatory hypotheses [and comments and reply]. *Current Anthropology*, *26*(2), 223–246. https://doi.org/10.1086/203251

Kaplan, H., Kim, H., & Hurtado, A. M. (1990). Risk, foraging, and food sharing among the Ache. In E. Cashdan (Ed.), *Risk and uncertainty in tribal and peasant economies* (pp. 107–143). Routledge. https://doi.org/10.4324/9780429304866

Mengarelli, F., Moretti, L., Faralla, V., Vindras, P., & Sirigu, A. (2014). Economic decisions for others: An exception to loss aversion law. *PLOS ONE*, *9*(1), e85042. https://doi.org/10.1371/journal.pone.0085042

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*(1), 1–10. https://doi.org/10.1038/s41598-018-31308-6

Polman, E. (2012). Self–other decision making and loss aversion. *Organizational Behavior and Human Decision Processes*, *119*(2), 141–150. https://doi.org/10.1016/j.obhdp.2012.06.005

Shenhav, A., & Greene, J. D. (2010). Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron*, *67*(4), 667–677. https://doi.org/10.1016/j.neuron.2010.07.020

Zhong, S., Chark, R., Hsu, M., & Chew, S. H. (2016). Computational substrates of social norm enforcement by unaffected third parties. *NeuroImage*, *129*, 95–104. https://doi.org/10.1016/j.neuroimage.2016.01.040