ISSN (Print) 0473-453X
Discussion Paper No. 1262
ISSN (Online) 2435-0982

DETERRING EXPLOITATION: SOCIAL NORMS AND PUNISHMENT TECHNOLOGY

Michalis Drouvelis
Nobuyuki Hanaki
Yuta Shimodaira

Secondly Revised June 2026
Revised May 2025
November 2024

The Institute of Social and Economic Research
The University of Osaka
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Deterring Exploitation: Social Norms and Punishment Technology*

Michalis Drouvelis[†] Nobuyuki Hanaki[‡] Yuta Shimodaira[§]

June 3, 2026

Abstract

When are actors deterred from exploiting those over whom they hold unilateral power? We study this question in a laboratory experiment using a 2×2 design that extends the power-to-take game, varying whether the proposer can give as well as take, and whether the responder can destroy the proposer's endowment at no cost (costless retaliation). Either variation alone leaves proposer behaviour unchanged, even though costless retaliation substantially increases punishment. Only when giving is feasible and retaliation is costless do proposers take significantly less, with average take rates falling from 60% to below 40%. Our findings show how institutional structures—through available action sets and punishment technologies—jointly determine whether exploitation is deterred.

Keywords: Power-to-take game, Giving, Retaliation, Negative reciprocity, Social norms

JEL Code: A12, C72, C91

---

*The experiment reported in this paper was approved by the Institutional Review Board of the Institute of Social and Economic Research (ISER), the University of Osaka (No. 20230801), and was pre-registered at AsPredicted (#80626, `https://aspredicted.org/d73jg.pdf`; and #179384, `https://aspredicted.org/kzsk-2c3p.pdf`). Authors have employed LLMs (OpenAI's ChatGPT and Anthropic's Claude) to proofread and polish the manuscript.

Abbreviations list: CLR, costless retaliation; G, giving; JPY, Japanese Yen; KS, Kolmogorov–Smirnov; KW, Kruskal–Wallis; MW, Mann–Whitney; PTG, power-to-take game; SD, standard deviation.

[†]University of Birmingham and CESifo, Munich. E-mail: m.drouvelis@bham.ac.uk

[‡]Corresponding author. Institute of Social and Economic Research, the University of Osaka, and University of Limassol. E-mail: nobuyuki.hanaki@iser.osaka-u.ac.jp

[§]Shinshu University. E-mail: yshimodaira@shinshu-u.ac.jp

1 Introduction

When does punishment deter exploitation? A large body of research shows that many individuals are willing to punish unfair behaviour, even when doing so is personally costly and yields no future material benefit (Fehr and Gächter, 2000, 2002; de Quervain et al., 2004; Drouvelis, 2021). The possibility of punishment can substantially alter behaviour, promoting cooperation and discouraging opportunistic actions (Fehr and Gächter, 2000; Gürerk et al., 2006). Yet punishment does not always prevent exploitation. Even when opportunities to punish exist, exploitative behaviour often persists, suggesting that the mere availability of punishment is insufficient to deter misconduct. Understanding the conditions under which punishment actually deters exploitation therefore remains an important open question.

Two features of the institutional environment are likely to be particularly important. First, the willingness to punish depends not only on outcomes but also on how actions are interpreted. The same outcome may be judged more harshly when the decision maker could have chosen a more generous alternative (Falk et al., 2003, 2008). Second, punishment may fail to deter exploitation if carrying it out is sufficiently costly that the threat lacks credibility (de Quervain et al., 2004). Understanding when punishment deters exploitation therefore requires considering both how actions are normatively evaluated and how punishment can be implemented.

This paper investigates how these two features of the institutional environment determine how people exploit others in a controlled laboratory experiment. In a 2×2 between-subjects design, we extend the power-to-take game (PTG) of Bosman and van Winden (2002) by independently varying two institutional features. First, we vary whether the proposer—the party with unilateral power over the responder's endowment—can either take from or give to the responder (the giving dimension, G). Second, we vary whether the responder can destroy the proposer's endowment at no personal cost (costless retaliation, CLR) or only their own endowment (costly retaliation, as in the standard PTG). This yields four treatments—PTG, PTG+G, PTG+CLR, and PTG+G+CLR—that allow us to isolate the separate and joint effects of counterfactual generosity induced by the possibility to give and punishment technology on proposers' willingness to exploit. More broadly, the design allows us to distinguish situations in which punishment is feasible from situations in which punishment is both feasible and targeted at behaviour that is perceived as normatively inappropriate. This distinction is central to understanding when punishment actually deters exploitation.

The PTG is well suited to our investigation because it provides continuous measures of both exploitation and retaliation (punishment). In its standard format, individuals in PTG are endowed with a fixed amount of money. The proposer chooses a *take rate* between $[0, 1]$ from the responder's endowment, and the responder, after observing this, chooses a *destruction rate* between $[0, 1]$, i.e., the fraction of own endowment to destroy. This continuity is essential as it allows us to observe not only by how much proposers' exploitation changes but also by how much the severity of responders' retaliation changes across treatments—nuances that the ultimatum game cannot capture.

The structure of our extended PTG maps naturally onto situations where one party holds unilateral authority over another's resources, while the affected party retains the ability to retaliate in non-contractible yet consequential ways. Consider a manager with discretion to reward or penalise employees. Employees may retaliate through actions that are personally costly or through actions that primarily harm the manager. The moral evaluation of the manager's decision depends on the available alternatives: penalising an employee when a reward was feasible is judged more harshly than penalising when no reward was available. Whether such judgments deter exploitation depends on the credibility of retaliation.

Our results are striking. Adding only the giving option (PTG+G) or only costless retaliation (PTG+CLR) does not significantly affect proposers' behaviour compared to the baseline PTG. The average take rate is about 60% in all three treatments. This is despite the fact that responders destroy proposers' endowments far more frequently in PTG+CLR than in PTG. Only in PTG+G+CLR where both conditions are present—responders can retaliate costlessly and proposers have the option to give—does proposer behaviour change substantially, with the average take rate falling below 40%.

We interpret these findings as follows: When retaliation is costly, proposers anticipate little or no retaliation. Thus, they take regardless of whether the giving option is available. However, taking when the give option is available is widely viewed as socially inappropriate (Krupka and Weber, 2013); thus, when retaliation is costless, proposers refrain from taking by anticipating harsh retaliation by the responder. In contrast, when the option to give is absent, taking is judged less inappropriate; proposers may therefore expect that taking will not trigger severe retaliation and continue to take from responders even when retaliation is costless.

Neither individual-level inequality aversion (measured using an incentivised multiple price list task) nor changes in self-reported emotional states predict the treatment effects. Both leave treatment effects unchanged and explain negligible individual-level variance. Using the tractable reciprocity framework of Cox et al. (2007), we further show that the estimated weight proposers place on the responder's payoff is similar across all four treatments, indicating that underlying social preferences do not change—what changes is the proposer's anticipation of how their action will be normatively interpreted and punished.

Our paper contributes to three literatures. First, we identify the joint institutional conditions under which punishment deters exploitation. Second, we show that the availability of giving opportunities affects the normative interpretation of taking in a setting with endogenous retaliation. Third, we qualify the relationship between costless punishment and deterrence by showing that punishment deters exploitation only when the behaviour is perceived as normatively inappropriate.

The rest of the paper is organised as follows. Section 2 outlines the experimental design and procedures. Section 3 presents our results, and Section 4 concludes.

2 Experimental Design

The main framework we consider in our experiment is a one-shot PTG, as introduced by Bosman and van Winden (2002). The basic structure involves a simple two-player game, in which each player is randomly assigned one of two roles: 'proposer' or 'responder' (called player 1 and player 2, respectively, in the instructions).

Participants first completed a real effort task based on Kamei and Markussen (2022) to determine their endowments.[1] The task was calibrated so that all participants earned the maximum endowment of 1,000 JPY.[2]

The real effort task is followed by the PTG. The game is sequential and consists of two stages. In Stage 1, the proposer chooses a fraction $t \in [0, 1]$ of the responder's endowment to take. In Stage 2, after observing the proposer's decision, t, the responder chooses a fraction $d \in [0, 1]$ of their endowment to destroy. Each player's earnings are given by:

- Proposer's earnings: $Y_1 + t(1 - d)Y_2$

[1]Participants face randomly generated sequences of ten binary digits (0s and 1s) and are asked to count the number of 1s.

[2]At the time of the experiment, 1 USD corresponded to approximately 114 JPY in 2021 and 160 JPY in 2024.

- Responder's earnings: $(1 - t)(1 - d)Y_2$

where Y_i is player i's endowment ($i = 1$ for the proposer and $i = 2$ for the responder). We set $Y_1 = Y_2 = 1{,}000$, since the real effort task was calibrated to yield an expected endowment of 1,000.

We extend this baseline framework (the PTG treatment) along two dimensions. Table 1 provides an overview. The first variation concerns whether proposers have the option to give part of their own endowment to the responder (the PTG+G and PTG+G+CLR treatments, if proposers can), in addition to taking part of the responder's endowment. The second concerns whether responders have the option to destroy part of the proposer's endowment (the PTG+CLR and PTG+G+CLR treatments, if responders can), in addition to destroying part of their own endowment in response to the proposer's decision. Thus, our experiment implements a 2×2 between-subjects design, resulting in four treatments (summarised in Table 1).

Let $g \in [0, 1]$ denote the fraction of their endowment that proposers give to the responder (for PTG+G and PTG+G+CLR), and let $r \in [0, 1]$ denote the fraction of the proposer's endowment that responders destroy (for PTG+CLR and PTG+G+CLR). Each player's earnings under the four treatments are also summarised in Table 1.

The decision-making interfaces are described as follows (see Online Appendix II.5 for instructions, comprehension quizzes, and screenshots of the decision-making screens). On the proposer's decision screen, both the proposer's and the responder's earnings from the preceding real effort task are

Table 1: Overview of experimental treatments

	Proposer can take part of responder's endowment	Proposer can take part of responder's endowment or give part of their own endowment to the responder
Responder can destroy their own endowment	PTG $\pi_1 = Y_1 + t(1-d)Y_2$ $\pi_2 = (1-t)(1-d)Y_2$	PTG with an option to give (PTG+G) $\pi_1 = (1-g)Y_1 + t(1-d)Y_2$ $\pi_2 = gY_1 + (1-t)(1-d)Y_2$ $(t \times g = 0)$
Responder can destroy either their own or the proposer's endowment	PTG with costless retaliation (PTG+CLR) $\pi_1 = (1-r)Y_1 + t(1-d)Y_2$ $\pi_2 = (1-t)(1-d)Y_2$ $(r \times d = 0)$	PTG+G with costless retaliation (PTG+G+CLR) $\pi_1 = (1-r)(1-g)Y_1 + t(1-d)Y_2$ $\pi_2 = (1-r)gY_1 + (1-t)(1-d)Y_2$ $(t \times g = 0,\ r \times d = 0)$

π_i: Payment to player i ($i = 1$ for the proposer; $i = 2$ for the responder).
$t \in [0, 1]$: Fraction of the responder's endowment that the proposer decides to take.
$d \in [0, 1]$: Fraction of the responder's own endowment that the responder decides to destroy.
$g \in [0, 1]$: Fraction of the proposer's endowment that the proposer gives to the responder (in PTG+G and PTG+G+CLR).
$r \in [0, 1]$: Fraction of the proposer's endowment that the responder destroys (in PTG+CLR and PTG+G+CLR).

displayed. In treatments where the proposer has a 'give' option (PTG+G and PTG+G+CLR), the proposer must choose either to take from or to give to the responder. After selecting one of these two options, they enter an integer percentage (0–100) corresponding to the amount taken or given.[3] In treatments without the give option (PTG and PTG+CLR), the proposer can only take from the responder and enters the corresponding percentage.

Once the proposer's decision is complete, control passes to the responder. On the responder's decision screen, both participants' earnings from the real effort task are shown alongside the proposer's choice (i.e., the percentage taken or given). In treatments where the responder can destroy the proposer's endowment (PTG+CLR and PTG+G+CLR), the responder must choose either to destroy part of their own endowment or part of the proposer's endowment. After selecting one of these two options, they enter an integer percentage (0–100) to destroy.[4] In treatments without this option (PTG and PTG+G), the responder can only destroy part of their own endowment and enters the corresponding percentage.

After the responder finalises their decision, both participants receive feed-

[3]When proposers with the 'give' option wished to refrain from both taking and giving (i.e., $t = g = 0$), they were required to choose one of the options and then set the percentage to zero. In our data analysis, when the percentage was zero, we did not distinguish which option had been selected.

[4]When responders with the 'destroy proposer's endowment' option wished to refrain from any destruction (i.e., $d = r = 0$), they were required to choose one of the options and then set the percentage to zero. In our data analysis, when the percentage was zero, we did not distinguish which option had been selected.

back on the choices made and on their resulting earnings.

2.1 Additional tasks

As decisions to take (or to give), as well as to retaliate, may be influenced by participants' degrees of inequality aversion, we measure these degrees. To this end, we extend the dictator game task using multiple price lists, as proposed by Yang et al. (2016) and He and Wu (2016), to estimate the degree of advantageous and disadvantageous inequality aversion in the model of Fehr and Schmidt (1999).[5] This task is incentivised and is conducted before the real effort task and the PTG; participants are informed of the results only after all tasks have been completed. See Online Appendix I for details.

Previous evidence from experiments on the PTG has shown that emotions can explain some aspects of participants' behaviour (see Bosman and van Winden, 2002; Bosman et al., 2005; Galeotti, 2015). Following Bosman and van Winden (2002), we elicit the intensity with which participants experience each of the following emotions using a seven-point Likert scale (1 = 'no emotion at all'; 7 = 'high intensity of the emotion'): irritation, anger, contempt, envy, jealousy, sadness, joy, shame, fear, and surprise (see Online Appendix II.2 for the screenshot). We measure these emotions three times

[5]Consistent with Yang et al. (2016), who stress the importance of separating strategic and reciprocal motives from pure inequality aversion, we directly assess participants' level of disadvantageous inequality aversion using straightforward choice menus, as in Yang et al. (2016) and He and Wu (2016), rather than inferring it indirectly through the ultimatum game, as in Blanco et al. (2011).

during our experiment: before participants receive any instructions, after they have received instructions for the PTG, and once more after they have participated in the game and received feedback. This design allows us to investigate how emotional intensity changes across experimental phases.

2.2 Implementation

The experiment was computerised using oTree (Chen et al., 2016) and conducted in the laboratory of the Institute of Social and Economic Research at the University of Osaka, in December 2021 and June–July 2024.[6] Participants were recruited from the subject pool managed by ORSEE (Greiner, 2015), which consists of students from the University of Osaka. A total of 328 students (224 males, 102 females, and 2 others) participated in the experiment.[7] The experiment lasted approximately 73 minutes (SD = 6.7), including instruction and payment time. Average earnings across all tasks were 2,031 JPY (SD = 695), including a fixed fee of 500 JPY.

Online Appendix III shows that our samples are well balanced in terms of gender composition and the measured degrees of inequality aversion across the four treatments. As noted in Section 2, we designed the real effort task so

[6]The experiment was conducted in two waves. The second wave was implemented to increase statistical power.

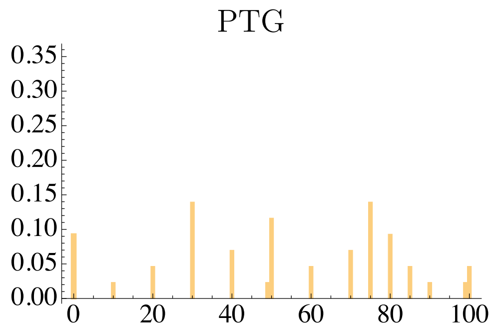
[7]We conducted four sessions for each treatment. The number of participants per treatment is as follows: PTG, 86; PTG+G, 80; PTG+CLR, 82; PTG+G+CLR, 80. Although we recruited the same number of participants for each session, the final sample sizes vary due to no-shows.

that every participant could earn the maximum endowment of 1,000 JPY for use in the PTG. Although participants needed to answer a minimum of 40 questions correctly, they actually answered substantially more. The results of the real effort task are reported in Online Appendix III.3.

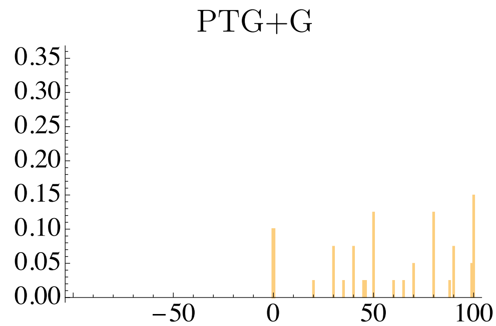
3 Results

Table 2 shows the average payoffs obtained by proposers and responders in the PTG across the four treatments. The bottom row reports the p-values from the Kruskal–Wallis (KW) test for each column. Proposers' payoffs are significantly lower in the treatments with costless retaliation (PTG+CLR and PTG+G+CLR) than in those without it (PTG and PTG+G) (Mann–Whitney (MW) test comparing groups with vs. without costless retaliation, $p < 0.001$). Responders' payoffs are higher in the treatments offering both the option to give and costless retaliation (PTG+G+CLR) than in the other three treatments (MW tests: $p = 0.070$ vs. PTG; $p = 0.023$ vs. PTG+CLR; $p = 0.014$ vs. PTG+G). Total payoffs per pair are lowest in PTG+CLR (MW tests: $p < 0.001$ vs. PTG; $p < 0.001$ vs. PTG+G; $p = 0.111$ vs. PTG+G+CLR). Below, we analyse behaviour of proposers and responders separately.

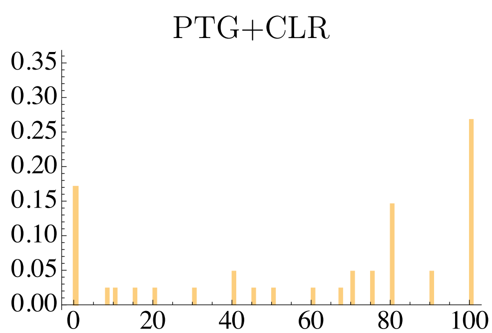
Figure 1: Distribution of proposer choices by treatment



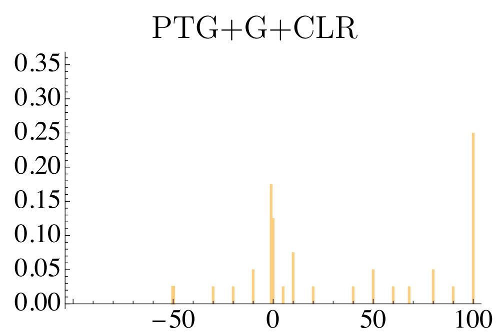
	Take	Total
N	39	43
mean	59.18	53.67
SD	(24.61)	(29.16)



	Give	Take	Total
N	0	36	40
mean	n.a.	67.42	60.68
SD	n.a.	(25.87)	(31.94)



	Take	Total
N	34	41
mean	71.62	59.39
SD	(29.52)	(38.25)



	Give	Take	Total
N	12	23	40
mean	10.58	68.39	36.15
SD	(15.56)	(36.22)	(47.55)

Note: The horizontal axis of each histogram represents the take rate (t) and the give rate (g), with t shown as a negative value. All values are expressed as percentages. In the table below the histogram, the 'Take' column includes observations with $t > 0$, the 'Give' column includes observations with $g > 0$, and the 'Total' column includes all observations, with giving coded as negative taking.

Table 2: Average payoffs by treatment

	Proposer	Responder	Total in pair
PTG	1449.01	457.97	1906.98
	(292.69)	(297.27)	(286.31)
PTG+CLR	950.98	406.1	1357.07
	(351.31)	(382.55)	(454.37)
PTG+G	1387.51	375.49	1763.0
	(325.93)	(328.66)	(412.52)
PTG+G+CLR	954.13	629.87	1584.0
	(326.95)	(466.07)	(452.86)
p-value in KW test	< 0.001	0.035	< 0.001

Note: Standard deviations are shown in parentheses.

3.1 Proposers' behaviour

Figure 1 shows the distribution of choices made by proposers in each treatment. In the treatments with an option to give (PTG+G and PTG+G+CLR), give rates are represented as negative take rates. Below the histogram, we report the number of proposers who chose a positive take rate and the number who chose a positive give rate, along with their respective average values. An eyeballing inspection of Figure 1 reveals that, when the give option is present and the responder cannot retaliate costlessly (i.e., in PTG+G), the proposer chooses not to give anything to the responder. Furthermore, the distributions of choices in PTG and PTG+G do not differ significantly ($p = 0.337$, Kolmogorov–Smirnov (KS) test).

The behaviour of proposers tends to be more extreme when costless retaliation is possible for responders. Comparing the two treatments with the option to give, the frequencies of 0% and 100% are clearly higher in

Table 3: Results of pairwise comparisons of proposer behaviour (p-values from MW tests)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	0.225	0.224	0.073
PTG+CLR	-	0.863	0.018
PTG+G	-	-	0.023

PTG+G+CLR than in PTG+G. Furthermore, in PTG+G+CLR, some proposers choose to give rather than take.[8] Similarly, compared with PTG, we observe more frequent take rates of 0% and 100% in PTG+CLR.

The overall average take rates (with giving coded as negative taking) by proposers across the four treatments are 53.67, 60.68, 59.39, and 36.15 for PTG, PTG+G, PTG+CLR, and PTG+G+CLR, respectively. Differences across treatments are significant ($p = 0.040$, KW test).[9] In particular, looking at proposers' overall behaviour, we find that they take, on average, significantly less in PTG+G+CLR compared with both PTG+G ($p = 0.023$, MW test) and PTG+CLR ($p = 0.018$, MW test). This indicates that the presence of both the give and retaliation options is important in leading proposers to take less (PTG+G+CLR vs. PTG: $p = 0.073$, MW test). By contrast, adding only the give option without retaliation (PTG vs. PTG+G:

[8] In Online Appendix IV.1, we report the results of probit regressions in which the dependent variable is a dummy equal to one if the proposer takes ($t > 0$) and zero otherwise. The results confirm that, in PTG+G+CLR, proposers are significantly less likely to take from responders than in the other three treatments.

[9] Any statistical tests reported in this section refer to two-sided tests unless otherwise stated.

$p = 0.224$, MW test) or adding only costless retaliation without giving (PTG vs. PTG+CLR: $p = 0.225$, MW test) does not significantly change behaviour relative to the baseline PTG treatment. Table 3 reports the pairwise MW tests.

Evidence from previous dictator games (e.g., List, 2007; Bardsley, 2008; Drouvelis, 2023) shows that allowing dictators to take money from their matched recipients reduces the amount allocated to the recipients. Unlike dictator games, where taking affects sharing, introducing giving options in the PTG treatment does not significantly affect behaviour.[10]

Let us further analyse the behaviour of proposers by examining the increase in their payoffs resulting from their own decisions ($tY_2 - gY_1$).[11] Column (1) in Panel (a) of Table 4 reports the coefficients from regressing $tY_2 - gY_1$ on the treatment dummies I_{+CLR}, I_{+G}, and I_{+G+CLR}, which correspond to PTG+CLR, PTG+G, and PTG+G+CLR, respectively. The baseline treatment is PTG. The estimated coefficients and the Wald tests in Panel (b) show that the increase in proposer's payoffs does not differ significantly across PTG, PTG+G, and PTG+CLR. However, it is significantly

[10]If we focus on proposers with positive take rates, the mean take rates are 59.18, 67.42, 71.62, and 68.39 for PTG, PTG+G, PTG+CLR, and PTG+G+CLR, respectively, and there is no significant difference across the four treatments ($p = 0.106$, KW test). Furthermore, Tobit regressions reported in Online Appendix IV.1 show no significant differences across treatments in taking behaviour once those who gave are excluded.

[11]Here, we focus on the stage prior to the responder's possible destruction of part of either endowment. Note that each proposer can either take or give, i.e., $t \times g = 0$. Since $Y_1 = Y_2$ in our experimental setup, the expression $tY_2 - gY_1$ is equivalent to the net take rate $t - g$.

Table 4: Treatment effects on payoff changes attributable to the proposer

(a) Regression results

	Increase in proposer's payoff due to proposer's decision $tY_2 - gY_1$	
	(1)	(2)
$I_{+\text{CLR}}$	57.158	64.442
	(81.348)	(82.466)
$I_{+\text{G}}$	70.006	65.141
	(81.867)	(83.188)
$I_{+\text{G+CLR}}$	−175.244**	−165.266*
	(81.867)	(84.474)
α		−132.310
		(331.289)
β		−33.555
		(43.633)
const.	536.744***	531.395***
	(56.833)	(57.988)
N	164	162
adj. R^2	0.047	0.041

Note: ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors are shown in parentheses.

(b) P-values for pairwise comparisons based on regression (1) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	0.483	0.394	0.034
PTG+CLR	-	0.877	0.006
PTG+G	-	-	0.004

lower in PTG+G+CLR than in the other three treatments, confirming our earlier non-parametric analyses.

In Column (2), we further control for the proposers' measured degrees of inequality aversion, α and β. Neither parameter is statistically significant, and the magnitude and significance of the treatment dummies remain similar to those reported in Column (1).

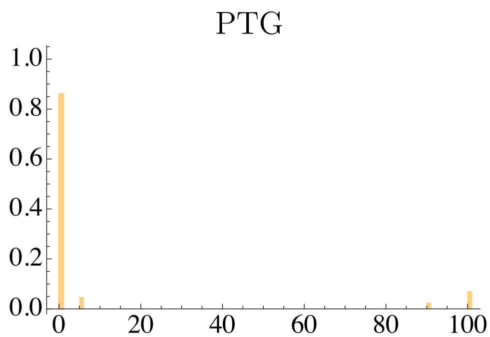
Finding 1 *On average, proposers take significantly less from responders only when both the giving option and costless retaliation are available.*

3.2 Responders' behaviour

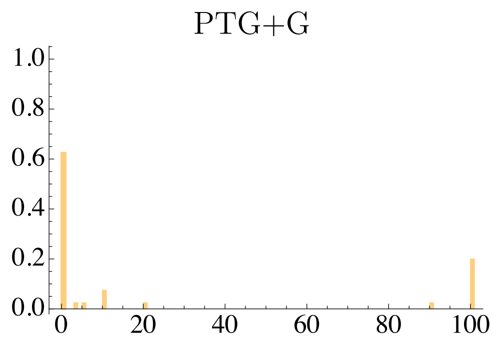
Let us now turn to the responders' behaviour. Figure 2 shows the distribution of choices made by responders in each treatment. In treatments with costless retaliation (PTG+CLR and PTG+G+CLR), retaliation rates are represented as negative destruction rates. Recall that the term 'destruction' refers to reductions in the responder's own endowment, whereas 'retaliation' refers to reductions in the proposer's endowment. Below each histogram, we report the number of responders who chose a positive destruction rate and the number who chose a positive retaliation rate, along with their respective average values.

The overall average destruction rates, with retaliation coded as negative destruction, were 9.30, 23.70, –54.54, and –35.00 in PTG, PTG+G, PTG+CLR, and PTG+G+CLR, respectively. These choices differ signifi-

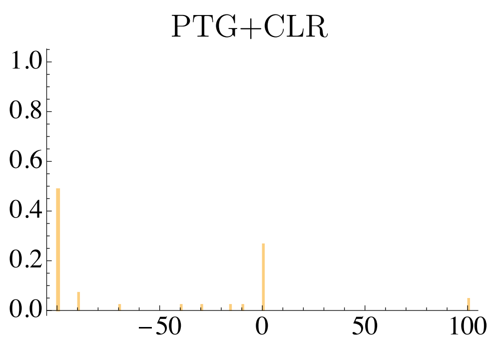
Figure 2: Distribution of responder choices by treatment



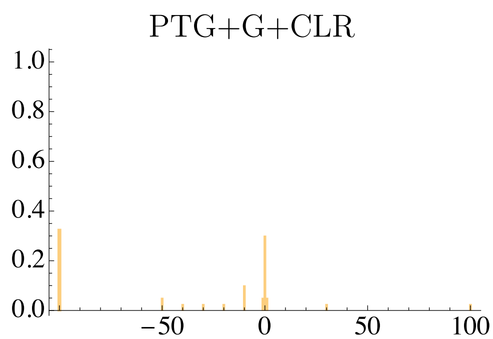
	Destroy	Total
N	6	43
mean	66.67	9.30
SD	(47.92)	(28.63)



	Destroy	Total
N	15	40
mean	63.20	23.70
SD	(45.45)	(41.25)



	Retaliate	Destroy	Total
N	28	2	41
mean	87.00	100	−54.54
SD	(27.30)	(0)	(57.03)



	Retaliate	Destroy	Total
N	24	4	40
mean	63.83	33.00	−35.00
SD	(41.97)	(46.71)	(50.69)

Note: The horizontal axis of each histogram represents the destruction rate of the responder's own endowment (d) and the retaliation rate of the proposer's endowment (r), with r shown as a negative value. All values are expressed as percentages. In the table below the histogram, the 'Destroy' column includes observations with $d > 0$, the 'Retaliate' column includes observations with $r > 0$, and the 'Total' column includes all observations, with retaliation coded as negative destruction.

Table 5: Results of pairwise comparisons of responder behaviour (p-values from MW tests)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	< 0.001	0.015	< 0.001
PTG+CLR	-	< 0.001	0.116
PTG+G	-	-	< 0.001

cantly across all four treatments ($p < 0.001$, KW test). When we compare the two treatments without costless retaliation (PTG vs. PTG+G), the difference remains significant ($p = 0.015$, MW test).[12]

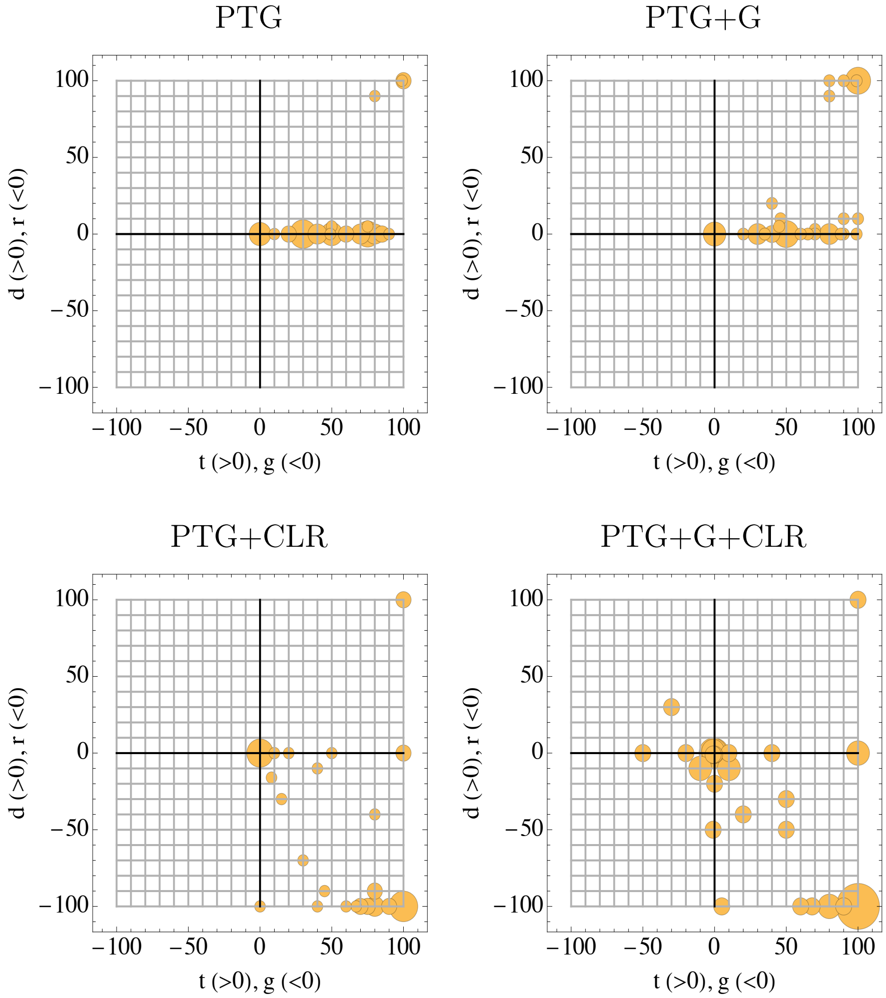
Comparing the two costless-retaliation treatments, the difference is not significant ($p = 0.116$, MW test), although proposers take less in PTG+G+CLR than in PTG+CLR. Table 5 summarises the p-values from all pairwise MW tests.

These analyses, however, do not account for the proposers' behaviour to which the responders reacted. We now turn to the joint behaviour of proposer–responder pairs. Figure 3 shows the joint distribution of proposer and responder choices across treatments.

In PTG+G, responders choose high destruction rates ($> 90\%$), even when proposers' take rates are below 100%. Such severe responses to take rates under 100% also occur in PTG, though much less frequently.

[12]Online Appendix IV.2 reports the results of a probit regression in which the dependent variable equals one if a responder chose to destroy part of their endowment ($d > 0$), and zero otherwise. The results show that responders are significantly more likely to destroy part of their endowment in PTG+G than in PTG.

Figure 3: Distribution of choice pairs by treatment



Note: Size of each bubble is proportional to the number of observations at its centre. The horizontal axis measures the take rate t (the give rate g is represented as a negative value), and the vertical axis measures the destruction rates d (the retaliation rate r is represented as a negative value).

In the treatments with costless retaliation, when proposers take 100% of the responder's endowment, responders either destroy their own endowment or that of the proposer; the payoff consequences are identical in both cases. A small number of responders opt not to react at all ($d = r = 0$) to a 100% take rate. By contrast, for take rates below 100%, many responders retaliate at a 100% rate (i.e., fully destroying the proposer's endowment) in both treatments.

To better understand responder behaviour, we analyse the decrease in both the proposer's and the responder's payoffs due to the responder's decision across all four treatments, while accounting for the intermediate payoff consequences of the proposer's choice.

Column (1) in Panel (a) of Table 6 reports the results of regressing the decrease in the proposer's payoff—relative to the intermediate payoff after the proposer's decision—due to the responder's decision[13] (i.e., $(1-g)rY_1+tdY_2$) on three treatment dummies (I_{+CLR}, I_{+G}, and I_{+G+CLR}, corresponding to PTG+CLR, PTG+G, and PTG+G+CLR, respectively). The baseline treatment is PTG. We focus on the decrease from the intermediate payoff because, as noted above, even with identical rates of retaliation or destruction, its impact in monetary terms depends on how much the proposer took from the responder.

Similarly, Column (4) in Panel (a) of Table 6 reports the results of regress-

[13] $[(1-g)Y_1 + tY_2] - [(1-g)(1-r)Y_1 + t(1-d)Y_2] = (1-g)rY_1 + tdY_2$.

Table 6: Treatment effects on payoff changes attributable to the responder

(a) Regression results

	Decrease in proposer's payoff due to responder's decision $(1-g)rY_1 + tdY_2$			Decrease in responder's payoff due to responder's decision $grY_1 + (1-t)dY_2$		
	(1)	(2)	(3)	(4)	(5)	(6)
$I_{+\text{CLR}}$	555.194***	515.275***	501.689***	−5.291	−5.059	−5.701
	(87.422)	(66.755)	(68.422)	(7.864)	(7.894)	(8.061)
$I_{+\text{G}}$	131.505	82.613	81.364	12.472	12.755	12.722
	(87.979)	(67.231)	(68.923)	(7.914)	(7.950)	(8.120)
$I_{+\text{G+CLR}}$	319.640***	442.031***	439.480***	3.337	2.627	1.773
	(87.979)	(68.031)	(69.499)	(7.914)	(8.045)	(8.188)
$t-g$		6.984***	7.077***		−0.041	−0.051
		(0.648)	(0.656)		(0.077)	(0.077)
α			−54.938			−7.215
			(253.143)			(29.825)
β			−29.306			−6.865**
			(24.911)			(2.935)
const.	87.733	−287.130***	−289.245***	5.291	7.465	8.454
	(61.076)	(58.114)	(59.562)	(5.494)	(6.872)	(7.017)
N	164	164	161	164	164	161
adj. R^2	0.206	0.538	0.538	0.013	0.009	0.033

Note: ***, ** ,*: statistically significantly different from zero at 1%, 5%, and 10% significance level. Standard errors are shown in parentheses.

Table 6: Treatment effects on payoff changes attributable to the responder (cont'd)

(b) P-values for pairwise comparisons based on regression (1) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	< 0.001	0.137	< 0.001
PTG+CLR	-	< 0.001	0.009
PTG+G	-	-	< 0.001

(c) P-values for pairwise comparisons based on regression (4) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	0.502	0.117	0.674
PTG+CLR	-	0.028	0.283
PTG+G	-	-	0.258

(d) P-values for pairwise comparisons based on regression (2) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	< 0.001	0.221	< 0.001
PTG+CLR	-	< 0.001	0.294
PTG+G	-	-	< 0.001

(e) P-values for pairwise comparisons based on regression (5) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	0.522	0.111	0.744
PTG+CLR	-	0.028	0.351
PTG+G	-	-	0.224

ing the decrease in the responder's payoff—relative to the same intermediate payoff—due to the responder's decision[14] (i.e., $grY_1 + (1-t)dY_2$) on the same three treatment dummies.

Column (1) shows that when costless retaliation is possible, the responder's decision reduces the proposer's payoffs significantly more in PTG+CLR

[14] $[gY_1 + (1-t)Y_2] - [g(1-r)Y_1 + (1-t)(1-d)Y_2] = grY_1 + (1-t)dY_2$.

than in PTG+G+CLR ($p = 0.009$, Wald test; see Panel (b)). Although this is consistent with the non-parametric analyses above, it arises because proposers take less in PTG+G+CLR than in PTG+CLR.

Column (4) shows that although the responder's decision reduces their own payoffs more in PTG+G than in PTG, the difference is not statistically significant. This result contradicts the non-parametric analyses above: when costless retaliation is not possible, responders are more likely to destroy their endowment if the proposer has the option to give than if they do not. The reduction in the responder's payoff is significantly greater in PTG+G than in PTG+CLR ($p = 0.028$, Wald test), but it is not significantly different from the other two treatments (see Panel (c)). This pattern persists when we control for proposer behaviour in Column (5) (see also Panel (e)).[15]

Columns (1) and (4) do not control for the proposer's behaviour. Because the responder's decision is likely influenced by the proposer's actions, Columns (2) and (5) include the proposer's net take rate ($t - g$) as a control. Columns (3) and (6) additionally control for the responder's inequality-aversion parameters, α and β. Across these specifications, a higher take rate by the proposer is associated with a significantly larger reduction in the proposer's payoff, but it does not significantly affect the reduction in the

[15]However, the Tobit regression reported in Online Appendix IV.2 shows that, if we restrict attention to responders who did not retaliate by destroying the proposer's endowment, the decrease in the responder's payoff due to their own behaviour is marginally significantly higher in PTG+G than in PTG.

responder's payoff.

Once we control for the proposer's behaviour in Columns (2) and (3), the difference in the reduction of the proposer's payoff between PTG+CLR and PTG+G+CLR is no longer significant ($p = 0.294$, Wald test; see Panel (d)). Both treatments still produce larger reductions than PTG and PTG+G.[16] Neither α nor β is significant in Column (3), suggesting that the degree of inequality aversion does not significantly predict the responder's decision to reduce the proposer's payoff.

However, Column (6) shows a negative and significant coefficient for β, indicating that the responder's degree of advantageous-inequality aversion is negatively correlated with the extent to which they reduce their own payoff. This finding is, however, difficult to interpret, primarily because the responders tend to occupy a disadvantageous position in our experiment.

Finding 2 *Controlling for the proposer's decision, when costless retaliation is possible, responders destroy the proposer's endowment regardless of whether the proposer has the option to give or not.*

Finding 3 *When costless retaliation is not possible, responders destroy their own endowment more frequently when proposers have the option to give than*

[16] This pattern persists when we restrict attention to responders who did not retaliate by destroying the proposer's endowment. Among these responders, the reduction in their own payoff under PTG+G+CLR is significantly larger than in the other three treatments. Thus, when proposers have the option to give, non-retaliating responders inflict greater destruction. See Online Appendix IV.2 for details.

when they do not; however, the effects on payoffs are not significantly different from those in the absence of that option.

To better understand the mechanisms underlying our results, Online Appendix IV.3 estimates the tractable reciprocity model of Cox et al. (2007).[17] The estimates suggest that the weight proposers place on responders' payoffs is broadly similar across treatments. Thus, the treatment effects appear to arise primarily from changes in anticipated retaliation and the normative interpretation of taking rather than from changes in underlying social preferences.

Furthermore, Online Appendix V presents analyses of emotional changes and their relationship to behaviour. We did not find any significant correlation between these emotional shifts and the responder's behaviour.

4 Conclusions

We study experimentally the conditions under which exploitation is deterred in a setting of unilateral power. Using a 2×2 design that extends the PTG

[17]This model is well-suited to our setting because it explicitly links observed destruction behaviour to an underlying 'emotional-status' parameter that captures the responder's reciprocal motivation. Our augmented PTG design—featuring both giving opportunities and variation in the costliness of retaliation—provides the exogenous variation required to identify separately how reciprocal motives respond to changes in the proposer's choice set and the responder's punishment technology. The model complements the reduced-form comparisons by quantifying how much of the observed heterogeneity in retaliation is driven by changes in underlying reciprocal preferences versus changes in beliefs or norms.

by varying (i) whether the proposer can give as well as take, and (ii) whether the responder can retaliate costlessly, we show that neither dimension alone significantly changes proposer behaviour, even though both increase responder punishment. Only when giving is possible and retaliation is costless do proposers take substantially less.

Our results indicate that neither giving opportunities nor costless retaliation alone substantially alter proposers' behaviour. Giving opportunities increase the perceived inappropriateness of taking, as reflected in responders' reactions, while costless retaliation increases the credibility of punishment. However, only when both features are present do proposers substantially reduce taking. Taken together, our findings suggest that exploitation is deterred not by the mere availability of punishment, but by punishment that is both credible and targeted at behaviour perceived as normatively inappropriate.

Future work should examine more directly the social norms and beliefs that underlie these responses and whether the same mechanisms operate in repeated interactions where retaliation may generate strategic incentives. Such work would help clarify when punishment can sustain cooperation and deter exploitation beyond the one-shot environments studied here.

Acknowledgment

We gratefully acknowledge financial support from the University of Birmingham, Japan Society for the Promotion of Science (JSPS) through Grant-in-Aid for Scientific Research (KAKENHI, grant numbers 18K19954, 20H05631, 23H00055, 25K16600, 25H00388, and 25K24681), as well as support from the Joint Usage/Research Center at ISER, the University of Osaka. We thank Ernst Fehr, Nicolaas Vriend, Daijiro Kawanaka, and seminar participants at the University of Osaka for their helpful comments. We also thank Natsumi Shimada, with whom we began this project. Unfortunately, she decided to withdraw from the project and declined co-authorship before the completion of this version of the paper.

References

Bardsley, N. (2008), 'Dictator game giving: altruism or artefact?', *Experimental Economics* **11**, 122–133.

Blanco, M., Engelmann, D. and Normann, H. T. (2011), 'A within-subject analysis of other-regarding preferences', *Games and Economic Behavior* **72**(2), 321–338.

Bosman, R., Sutter, M. and van Winden, F. (2005), 'The impact of real effort and emotions in the power-to-take game', *Journal of Economic Psychology* **26**, 407–429.

Bosman, R. and van Winden, F. (2002), 'Emotional hazard in a power-to-take experiment', *Economic Journal* **112**(476), 147–169.

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

Cox, J. C., Friedman, D. and Gjerstad, S. (2007), 'A tractable model of reciprocity and fairness', *Games and Economic Behavior* **59**, 17–45.

de Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A. and Fehr, E. (2004), 'The Neural Basis of Altruistic Punishment', *Science* **305**(5688), 1254–1258.

Drouvelis, M. (2021), *Social Preferences: An Introduction to Behavioural Economics and Experimental Research*, Agenda Publishing, Newcastle.

Drouvelis, M. (2023), 'Dictator giving and taking: Evidence from africa', *Journal of Economic Behavior and Organization* **209**, 326–333.

Falk, A., Fehr, E. and Fischbacher, U. (2003), 'On the nature of fair behavior', *Economic Inquiry* **41**(1), 20–26.

Falk, A., Fehr, E. and Fischbacher, U. (2008), 'Testing theories of fairness—intentions matter', *Games and Economic Behavior* **62**(1), 287–303.

Fehr, E. and Gächter, S. (2000), 'Cooperation and punishment in public goods experiments', *American Economic Review* **90**(4), 980–994.

Fehr, E. and Gächter, S. (2002), 'Altruistic punishment in humans', *Nature* **415**(6868), 137–140.

Fehr, E. and Schmidt, K. M. (1999), 'A theory of fairness, competition, and cooperation', *The Quarterly Journal of Economics* **114**, 817–868.

Galeotti, F. (2015), 'Do negative emotions explain punishment in power-to-take game experiments?', *Journal of Economic Psychology* **49**, 1–14.

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

Gürerk, Ö., Irlenbusch, B. and Rockenbach, B. (2006), 'The competitive advantage of sanctioning institutions', *Science* **312**(5770), 108–111.

He, H. and Wu, K. (2016), 'Choice set, relative income, and inequality aversion: An experimental investigation', *Journal of Economic Psychology* **54**, 177–193.

Kamei, K. and Markussen, T. (2022), 'Free Riding and Workplace Democracy—Heterogeneous Task Preferences and Sorting', *Management Science* **69**(7), 3884–3904.

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

List, J. A. (2007), ‘On the interpretation of giving in dictator games’, *Journal of Political Economy* **115**(3), 482–493.

Yang, Y., Onderstal, S. and Schram, A. (2016), ‘Inequity aversion revisited’, *Journal of Economic Psychology* **54**, 1–16.

Online Appendix for "Deterring Exploitation: Social Norms and Punishment Technology"

I Measuring inequality aversion

We extend the dictator game task using multiple price lists, as proposed by Yang et al. (2016) and He and Wu (2016), to estimate the degree of advantageous and disadvantageous inequality aversion in the model of Fehr and Schmidt (1999). In this task, each participant makes a series of binary choices between options A and B that determine payoffs for themselves and for another participant. The choice problems are presented in two lists: the first measures the degree of disadvantageous inequality aversion (α), and the second measures advantageous inequality aversion (β) (see Online Appendix II.3 for instructions, comprehension quizzes, and screenshots of the decision-making screens).

Each list contains 31 binary-choice questions, and all monetary values are expressed in JPY. The first list begins with a choice between option A (600 for you, 600 for the other) and option B (400 for you, 1,040 for the other), and ends with a choice between option A (0 for you, 600 for the other) and option B (400 for you, 1,040 for the other). The second list begins with a choice between option A (1,407 for you, 360 for the other) and option B (1,267 for you, 200 for the other), and ends with a choice between option A (360 for you, 360 for the other) and option B (1,267 for you, 200 for the other).

We based our price lists on those used by He and Wu (2016). However, in their task, the maximum measurable degree of advantageous inequality aversion, β, was only about 0.67. To allow for the possibility that some participants might exhibit higher β values, we extended the second list so that the maximum β was 0.85.

In each list, option B is constant, while the amount offered in option A decreases as one moves down the list. The row at which a participant switches from A to B therefore identifies the participant's degree of disadvantageous and advantageous inequality aversion. In our experiment, we impose a single switching point per list—ensuring consistency in decision-making—a procedure repeatedly used in previous studies (see Gonzalez and Wu, 1999; Jacobson and Petrie, 2009; Csermely and Rabas, 2016). Specifically, participants were presented with a price list of 31 rows, each presenting two radio-button options. To guarantee a single switch, the interface was programmed so that, once a button in a given row was selected, all rows above it were automatically assigned to option A and all rows below it to option B.[1]

This dictator game, used to elicit inequality aversion, was incentivised, as was the power-to-take game. Specifically, two participants were randomly paired; after both members of a pair had completed their responses, one of them was randomly selected as the 'dictator', and participants were paid based on the dictator's choice in one randomly selected item. Which of the 62 binary choices was selected was determined for each pair.

To prevent behaviour in the dictator game from influencing behaviour in the PTG, pairings were reassigned independently, and participants were informed of this in advance. In addition, the dictator game was conducted before the real effort task and the PTG, and the earnings from the dictator game were fed back to participants only at the very end, after all tasks had been completed.

[1] He and Wu (2016) had participants record their switch points on a paper decision sheet, which required detailed instructions. In our fully computerised experiment, we employ an interactive interface that presents choices so that shifts between alternatives are immediately and intuitively clear to participants.

II English translation of the instructions and screenshots

The instructions and screens have been translated from the original Japanese. All materials other than the window screenshots were printed and distributed to the participants. The descriptions shown in red are annotations.

II.1 Introduction

Thank you for your participation in our experiment today. Please read the following instructions carefully.

II.1.1 General instructions

- Please follow the experimenter's instructions.
- Only perform the operations specified.
- Do not talk to or exchange notes with other participants during the experiment.
- Do not attempt to peek at other participants' screens.
- Turn off your cell phone and place it in your bag.
- If you have any questions, quietly raise your hand to let the experimenter know.

II.1.2 Flow of the experiment

- Today's experiment will proceed as follows. It will take approximately 90 minutes.

  1. Survey
  2. Experiment 1: instructions, quiz, and task
  3. Experiment 2: instructions, practice session, and task
  4. Experiment 3: instructions and quiz
  5. Survey
  6. Experiment 3: task
  7. Survey

8. Payment

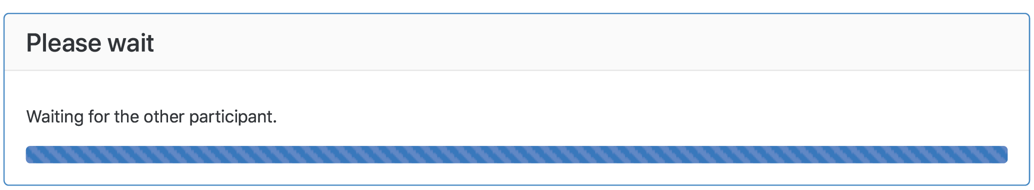
- Each experiment will be explained just before it begins. Please make sure you understand the instructions.
- There are two types of experiments: individual tasks and paired tasks. In paired tasks, you will be matched with another participant; pairings will be randomly reassigned each time.
- After we explain the earnings structure and how to operate the computer, we will proceed to Experiment 1.

II.1.3 Payment

- You will receive a participation fee of ¥500.
- In addition to the participation fee, you will receive the total amount you earn across the three experiments.
- Your total payment will be rounded up to the nearest ¥10.

II.1.4 Operation of the computer

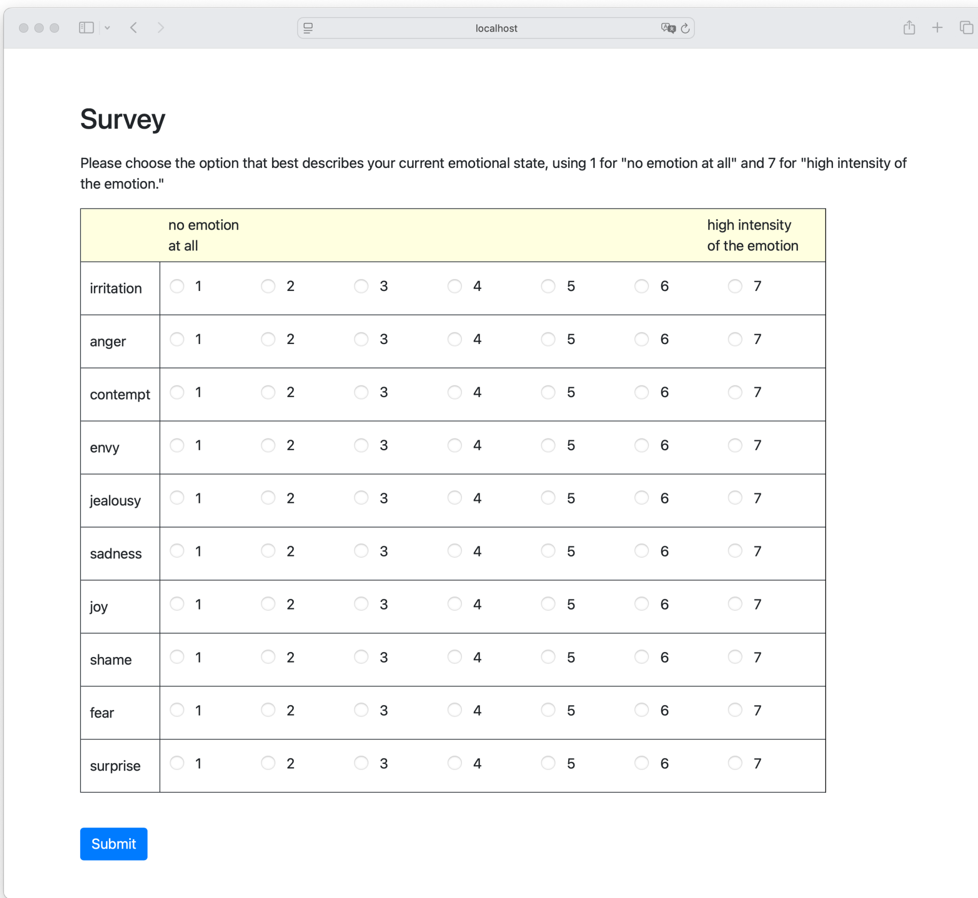
- Only the mouse and keyboard will be used in this experiment.
- Click the ‘Next’ button to proceed to the decision screen.
- On each decision screen, click the ‘Submit’ button to submit your decision.
- You will not be able to return to the previous screen.
- When your screen appears as shown below, your partner is making their decision. Please wait. After your partner has made their decision, the screen will update.



II.2 Survey on emotional state before the instructions for Experiment 1

[The survey on emotional state was conducted at three points, all on the same decision screen: 1) before the instructions for Experiment 1, 2) after the comprehension quiz for Experiment 3, and 3) after the feedback on the results of Experiment 3.]

Screenshot



localhost

Survey

Please choose the option that best describes your current emotional state, using 1 for "no emotion at all" and 7 for "high intensity of the emotion."

	no emotion at all						high intensity of the emotion
irritation	1	2	3	4	5	6	7
anger	1	2	3	4	5	6	7
contempt	1	2	3	4	5	6	7
envy	1	2	3	4	5	6	7
jealousy	1	2	3	4	5	6	7
sadness	1	2	3	4	5	6	7
joy	1	2	3	4	5	6	7
shame	1	2	3	4	5	6	7
fear	1	2	3	4	5	6	7
surprise	1	2	3	4	5	6	7

Submit

II.3 Experiment 1 (Inequality-aversion measurement task)

In Experiment 1, you will be randomly paired with another participant. Your earnings will be determined by either your decision or your partner's.

Before proceeding to the main experiment, you must complete a quiz to confirm your understanding of the experimental setup.

II.3.1 Decision-making task

- Two tables will appear on the screen. You can also find them on the next page [see Figure II.1].
- One of the tables contains 31 questions.
- For each question, you must choose either option A or option B.
- Each option specifies how the earnings are divided between you and your partner. For example, in the first row of the left-hand table, option A gives you ¥600 and your partner ¥600, while option B gives you ¥400 and your partner ¥1,040.
- The allocation for option B is the same for every question. In option A, the amount allotted to you decreases as you move down the table.
- Many participants choose option A for the first few questions and then switch to option B. To simplify the procedure, once you select option B for a given question, all questions in rows above it will automatically be set to option A, and all questions in rows below it will be set to option B.

II.3.2 Earnings

- Your earnings in Experiment 1 will be determined as follows. The exact amount will be revealed once you have completed all of today's experimental tasks.
- After both members of each pair have finalised their decisions, the computer will randomly select one of the 62 questions from the two tables for that pair. It will then randomly decide whether to use your decision or your partner's decision to determine your earnings.

If your decision is implemented For example, suppose that you chose option A, as shown in the figure below.

Option A				Option B	
Yours	Other's			Yours	Other's
200	600	◉	○	400	1040

If this question is selected and your decision is implemented, you will receive ¥200 and the other person will receive ¥600.

If your partner's decision is implemented For example, suppose your partner chose option B for the same question, as shown in the figure below.

Option A				Option B	
Yours	Other's			Yours	Other's
200	600	○	◉	400	1040

If this question is selected and your partner's decision is implemented, you will receive ¥1,040 and your partner will receive ¥400.

1st

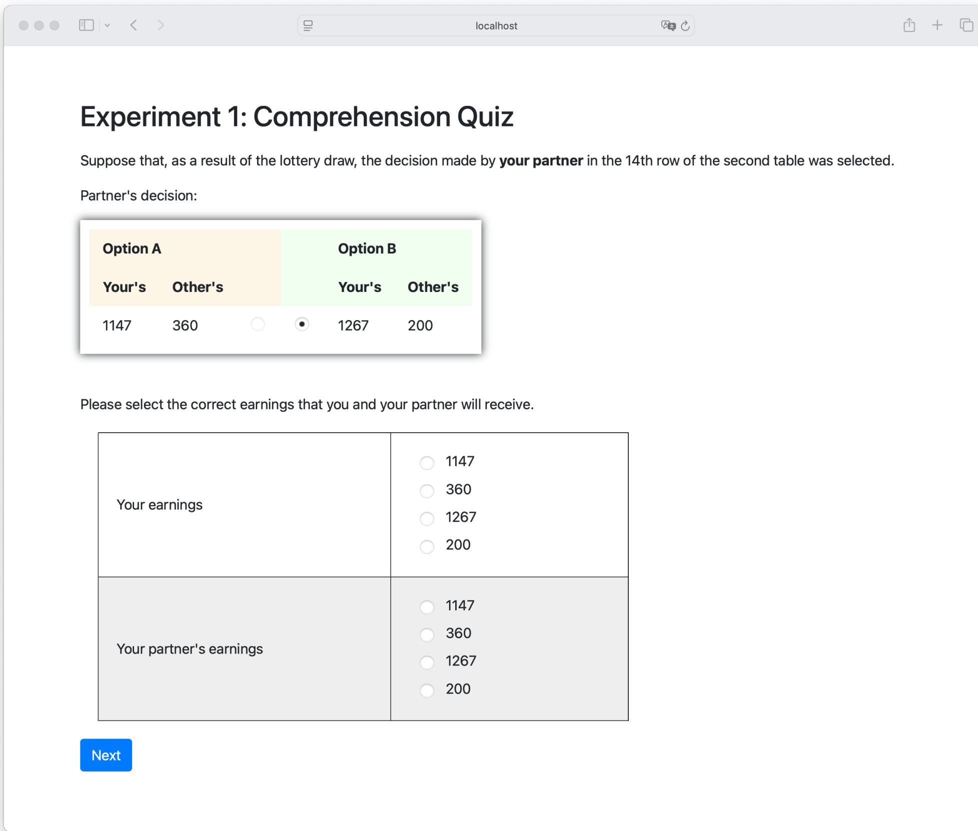
Option A		Option B	
Yours	Other's	Yours	Other's
600	600	400	1040
580	600	400	1040
560	600	400	1040
540	600	400	1040
520	600	400	1040
500	600	400	1040
480	600	400	1040
460	600	400	1040
440	600	400	1040
420	600	400	1040
400	600	400	1040
380	600	400	1040
360	600	400	1040
340	600	400	1040
320	600	400	1040
300	600	400	1040
280	600	400	1040
260	600	400	1040
240	600	400	1040
220	600	400	1040
200	600	400	1040
180	600	400	1040
160	600	400	1040
140	600	400	1040
120	600	400	1040
100	600	400	1040
80	600	400	1040
60	600	400	1040
40	600	400	1040
20	600	400	1040
0	600	400	1040

2nd

Option A		Option B	
Yours	Other's	Yours	Other's
1407	360	1267	200
1387	360	1267	200
1367	360	1267	200
1347	360	1267	200
1327	360	1267	200
1307	360	1267	200
1287	360	1267	200
1267	360	1267	200
1247	360	1267	200
1227	360	1267	200
1207	360	1267	200
1187	360	1267	200
1167	360	1267	200
1147	360	1267	200
1127	360	1267	200
1107	360	1267	200
1087	360	1267	200
1067	360	1267	200
1047	360	1267	200
1027	360	1267	200
1007	360	1267	200
987	360	1267	200
967	360	1267	200
947	360	1267	200
911	360	1267	200
868	360	1267	200
814	360	1267	200
747	360	1267	200
659	360	1267	200
538	360	1267	200
360	360	1267	200

Figure II.1: Two multiple price lists used in Experiment 1

Screenshot of the comprehension quiz



localhost

Experiment 1: Comprehension Quiz

Suppose that, as a result of the lottery draw, the decision made by **your partner** in the 14th row of the second table was selected.

Partner's decision:

Option A			**Option B**	
Your's	**Other's**		**Your's**	**Other's**
1147	360	○ ◉	1267	200

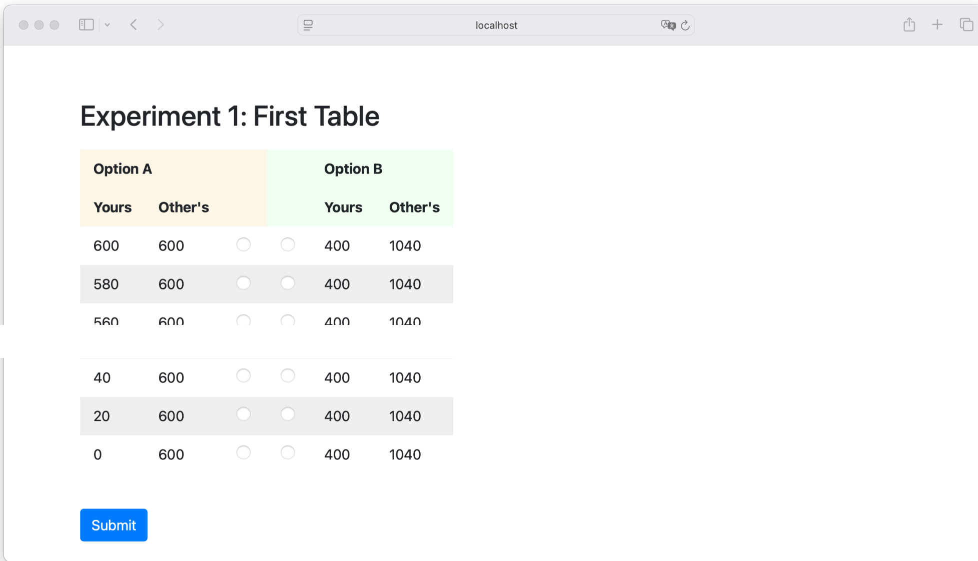
Please select the correct earnings that you and your partner will receive.

Your earnings	○ 1147 ○ 360 ○ 1267 ○ 200
Your partner's earnings	○ 1147 ○ 360 ○ 1267 ○ 200

Next

[Correct answers to the comprehension quiz: Your earnings are 200, and your partner's earnings are 1267.]

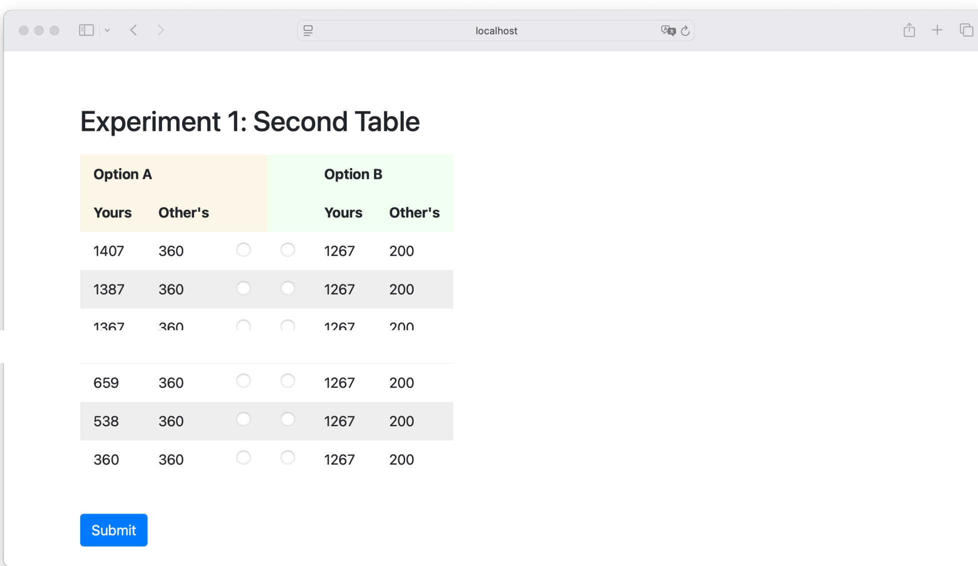
Screenshots of the decision screen


localhost

Experiment 1: First Table

Option A				Option B	
Yours	Other's			Yours	Other's
600	600			400	1040
580	600			400	1040
560	600			400	1040
40	600			400	1040
20	600			400	1040
0	600			400	1040

Submit


localhost

Experiment 1: Second Table

Option A				Option B	
Yours	Other's			Yours	Other's
1407	360			1267	200
1387	360			1267	200
1367	360			1267	200
659	360			1267	200
538	360			1267	200
360	360			1267	200

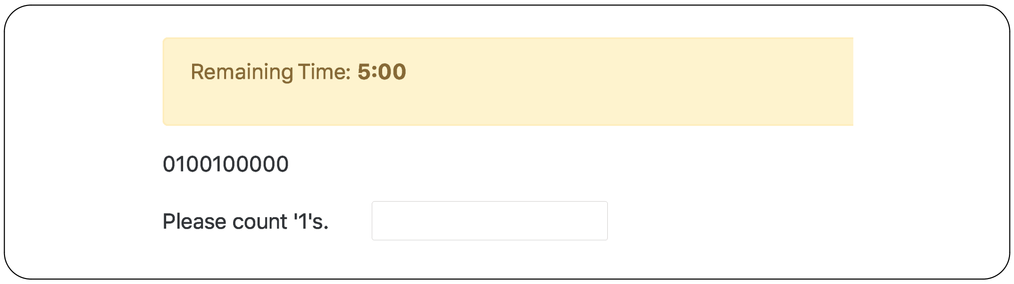
Submit

[Note that once the task had been completed, no feedback on the results was provided to the participants, and the procedure moved immediately on to Experiment 2.]

II.4 Experiment 2 (Real effort task)

II.4.1 Task

- On the screen, you will see a random 10-digit sequence of 0s and 1s, as as shown below.



- Count the number of 1s and enter your answer.
- In the example, the correct answer is 2.
- You have a time limit of 5 minutes. Answer as many sequences as you can within that time.

II.4.2 How to operate

- Use the numeric keypad or number keys to enter your response, then press 'Enter' key to submit.
- First, complete a 2-minute practice session, then proceed to the 5-minute main session.
- When practice ends, a 'Start' button will appear. Press it to begin the main session whenever you are ready.
- If you have any questions, quietly raise your hand and call the experimenter before starting the main session.

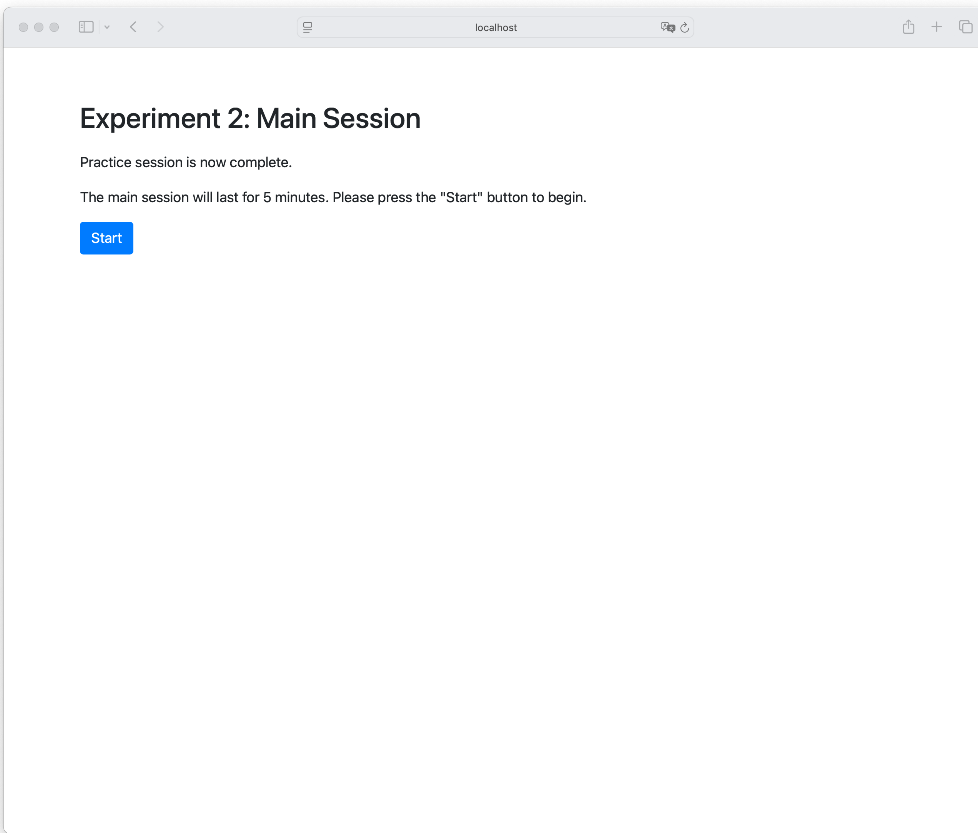
II.4.3 Earnings

- Only correct answers from the main session count toward your earnings in Experiment 2. Answers from the practice session are excluded.
- You will earn ¥25 for each correct answer. If your total number of correct answers meets or exceeds a predetermined threshold, your earnings will be ¥1,000.

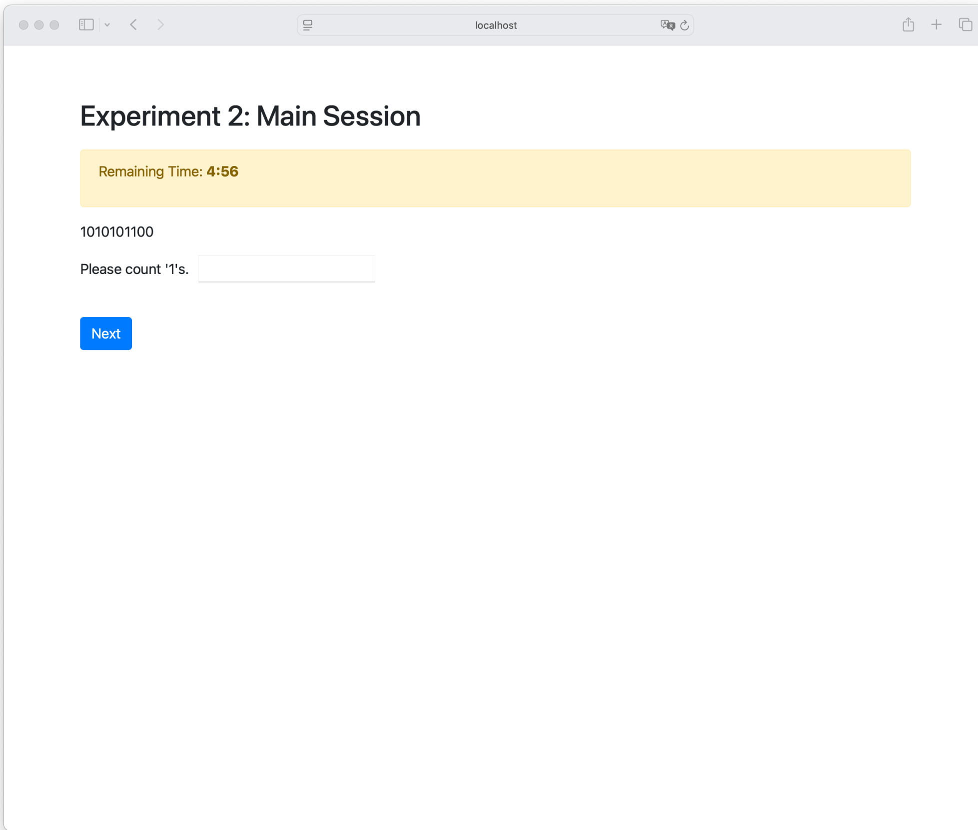
- In Experiment 3, which follows, you will perform a decision-making task using the amount you earned in Experiment 2.

Screenshots of the the main session

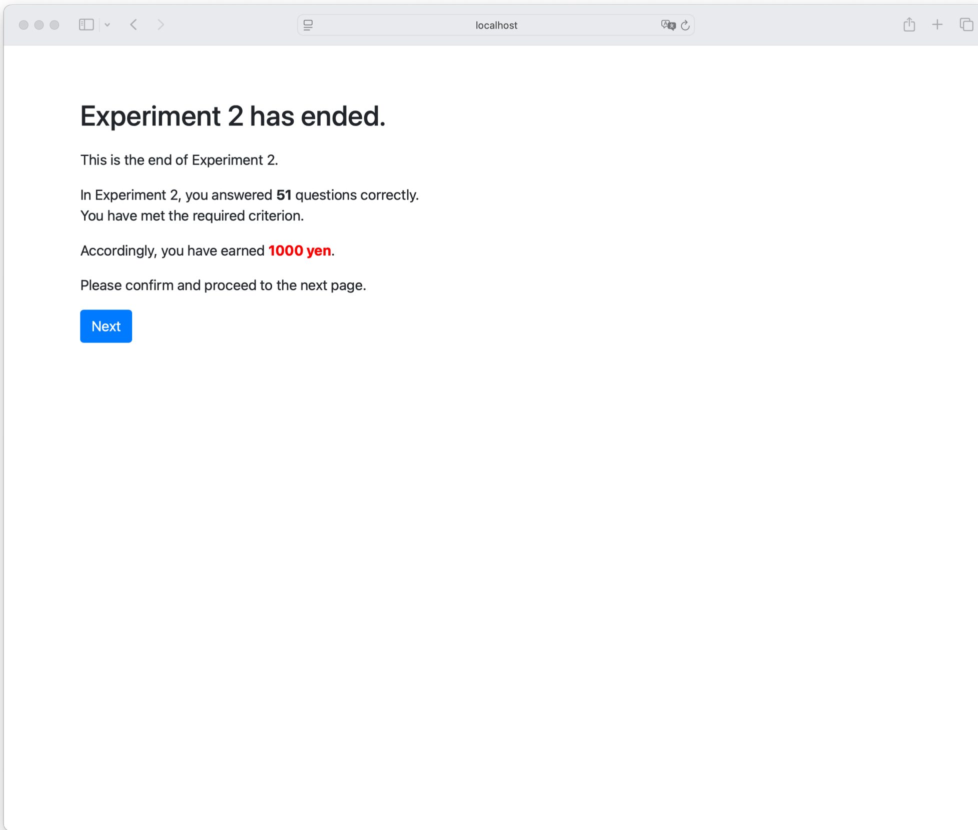
[The screen displayed after participants complete the practice session and before the main session begins]



[The screen of the the main session]



[The feedback screen for the results of the main session]



II.5 Experiment 3 (The power-to-take game)

In Experiment 3, you will be randomly paired with another participant. Each member of a pair will be assigned the role of Player 1 or Player 2, and you will make a one-time decision in that role.

Experiment 3 will uses the earnings you accumulated in Experiment 2.

Before proceeding to the main experiment, you must complete a quiz to confirm your understanding of the experimental setup and fill out a survey.

II.5.1 Decision-making task

First, Player 1 makes a decision. Then, Player 2 observes the outcome and makes a decision.

If you are Player 1,

[For PTG and PTG+CLR]
you can take the amount Player 2 earned in Experiment 2. Please specify the percentage you wish to take.

[For PTG+G and PTG+G+CLR]
you can choose either:

- to take the amount Player 2 earned in Experiment 2, or
- to give the amount you earned in Experiment 2.

Please choose one option and specify the percentage you wish to take or give.

If you are Player 2, after seeing Player 1's decision,

[For PTG and PTG+G]
you can destroy the amount you earned in Experiment 2. Please specify the percentage you wish to destroy.

[For PTG+CLR and PTG+G+CLR]
you can choose either:

- to destroy the amount Player 1 earned in Experiment 2, or

- to destroy the amount you earned in Experiment 2.

Please choose one option and specify the percentage you wish to destroy.

II.5.2 Earnings

Final earnings in Experiments 2 and 3 are determined by each player's decisions as follows. We will explain how your earnings are calculated in four cases, each based on a combination of decisions made by Player 1 and Player 2.

Case 1 [for PTG+CLR and PTG+G+CLR]

- Player 1 chooses to take the endowment that Player 2 earned in Experiment 2.
- Player 2 chooses to destroy the endowment that Player 1 earned in Experiment 2.

In this case:

- Player 1's earnings are equal to the portion of their endowment that Player 2 did not destroy, plus the amount they took from Player 2.
- Player 2's earnings are equal to their endowment minus the amount taken by Player 1.
- Specifically, the earnings are calculated using the following formulas:

$$\boxed{\text{Player 1's earnings}}$$
$$= \boxed{\begin{array}{c}\text{Player 1's earning}\\ \text{in Experiment 2}\end{array}} \times \left(100\% - \boxed{\begin{array}{c}\text{Fraction destroyed}\\ \text{by Player 2}\end{array}}\right)$$
$$+ \boxed{\begin{array}{c}\text{Player 2's earning}\\ \text{in Experiment 2}\end{array}} \times \boxed{\begin{array}{c}\text{Fraction taken}\\ \text{by Player 1}\end{array}}$$

$$\boxed{\text{Player 2's earnings}}$$
$$= \boxed{\begin{array}{c}\text{Player 2's earning}\\ \text{in Experiment 2}\end{array}}$$
$$- \boxed{\begin{array}{c}\text{Player 2's earning}\\ \text{in Experiment 2}\end{array}} \times \boxed{\begin{array}{c}\text{Fraction taken}\\ \text{by Player 1}\end{array}}$$

Example. Assume that the earnings in Experiment 2 are as follows:

- Player 1's earnings in Experiment 2: ¥1,000
- Player 2's earnings in Experiment 2: ¥1,000

First, Player 1 decides what percentage of Player 2's endowment to take. Suppose Player 1 decides to take 60%.

$$\boxed{\text{Fraction taken by Player 1}} = 60\%$$

Second, Player 2 reviews the percentage taken by Player 1 and decides what percentage of Player 1's endowment to destroy. Suppose Player 2 chooses to destroy 50%.

$$\boxed{\text{Fraction destroyed by Player 2}} = 50\%$$

The final amounts earned in Experiments 2 and 3 are as follows:

$$\begin{aligned} \text{Player 1's earnings} &= ¥1{,}100 \\ &= 1000 \times (100\% - 50\%) \qquad + 1000 \times 60\% \\ \text{Player 2's earnings} &= ¥400 \\ &= 1000 \qquad\qquad\qquad\qquad - 1000 \times 60\% \end{aligned}$$

Case 2 [for all treatments]

- Player 1 chooses to take the endowment that Player 2 earned in Experiment 2.
- Player 2 chooses to destroy the endowment that Player 2 earned in Experiment 2.

In this case:

- Player 1 may take from Player 2 any portion of Player 2's endowment that Player 2 did not destroy.
- Player 1's earnings are equal to their endowment plus the amount they took from Player 2.

- Player 2's earnings are equal to the portion of their endowment that they did not destroy, minus the amount taken by Player 1.
- Specifically, the earnings are calculated using the following formulas:

$$\boxed{\text{Player 1's earnings}}$$
$$= \boxed{\begin{array}{c}\text{Player 1's earning}\\ \text{in Experiment 2}\end{array}} + \boxed{\begin{array}{c}\text{Player 2's earning}\\ \text{in Experiment 2}\end{array}} \times \left(100\% - \boxed{\begin{array}{c}\text{Fraction destroyed}\\ \text{by Player 2}\end{array}}\right) \times \boxed{\begin{array}{c}\text{Fraction taken}\\ \text{by Player 1}\end{array}}$$

$$\boxed{\text{Player 2's earnings}}$$
$$= \boxed{\begin{array}{c}\text{Player 2's earning}\\ \text{in Experiment 2}\end{array}} \times \left(100\% - \boxed{\begin{array}{c}\text{Fraction destroyed}\\ \text{by Player 2}\end{array}}\right) - \boxed{\begin{array}{c}\text{Player 2's earning}\\ \text{in Experiment 2}\end{array}} \times \left(100\% - \boxed{\begin{array}{c}\text{Fraction destroyed}\\ \text{by Player 2}\end{array}}\right) \times \boxed{\begin{array}{c}\text{Fraction taken}\\ \text{by Player 1}\end{array}}$$

Example. Assume that the earnings in Experiment 2 are as follows:

- Player 1's earnings in Experiment 2: ¥1,000
- Player 2's earnings in Experiment 2: ¥1,000

First, Player 1 decides what percentage of Player 2's endowment to take. Suppose Player 1 decides to take 60%.

$$\boxed{\begin{array}{c}\text{Fraction taken}\\ \text{by Player 1}\end{array}} = 60\%$$

Second, Player 2 reviews the percentage taken by Player 1 and decides what percentage of Player 2's endowment to destroy. Suppose Player 2 chooses to destroy 50%.

$$\boxed{\begin{array}{c}\text{Fraction destroyed}\\ \text{by Player 2}\end{array}} = 50\%$$

The final amounts earned in Experiments 2 and 3 are as follows:

$$\begin{aligned}
\text{Player 1's earnings} &= \text{¥1,300} \\
&= 1000 \qquad\qquad\qquad\qquad + 1000 \times (100\% - 50\%) \times 60\% \\
\text{Player 2's earnings} &= \text{¥200} \\
&= 1000 \times (100\% - 50\%) \quad - 1000 \times (100\% - 50\%) \times 60\%
\end{aligned}$$

Case 3 [for PTG+G+CLR]

- Player 1 chooses to give the endowment that Player 1 earned in Experiment 2.
- Player 2 chooses to destroy the endowment that Player 1 earned in Experiment 2.

In this case:

- Player 1 may give to Player 2 any portion of Player 1's endowment that Player 2 did not destroy.
- Player 1's earnings are equal to the portion of their endowment that Player 2 did not destroy, minus the amount they gave to Player 2.
- Player 2's earnings are equal to their endowment plus the amount Player 1 gave to them.
- Specifically, the earnings are calculated using the following formulas:

$$\begin{aligned}
&\boxed{\text{Player 1's earnings}} \\
&= \boxed{\begin{matrix}\text{Player 1's earning}\\\text{in Experiment 2}\end{matrix}} \times \left(100\% - \boxed{\begin{matrix}\text{Fraction destroyed}\\\text{by Player 2}\end{matrix}}\right) \\
&\quad - \boxed{\begin{matrix}\text{Player 1's earning}\\\text{in Experiment 2}\end{matrix}} \times \left(100\% - \boxed{\begin{matrix}\text{Fraction destroyed}\\\text{by Player 2}\end{matrix}}\right) \times \boxed{\begin{matrix}\text{Fraction given}\\\text{by Player 1}\end{matrix}}
\end{aligned}$$

$$\begin{aligned}
&\boxed{\text{Player 2's earnings}} \\
&= \boxed{\begin{matrix}\text{Player 2's earning}\\\text{in Experiment 2}\end{matrix}} \\
&\quad + \boxed{\begin{matrix}\text{Player 1's earning}\\\text{in Experiment 2}\end{matrix}} \times \left(100\% - \boxed{\begin{matrix}\text{Fraction destroyed}\\\text{by Player 2}\end{matrix}}\right) \times \boxed{\begin{matrix}\text{Fraction given}\\\text{by Player 1}\end{matrix}}
\end{aligned}$$

Example. Assume that the earnings in Experiment 2 are as follows:

- Player 1's earnings in Experiment 2: ¥1,000
- Player 2's earnings in Experiment 2: ¥1,000

First, Player 1 decides what percentage of Player 1's endowment to give. Suppose Player 1 decides to give 60%.

$$\boxed{\text{Fraction given by Player 1}} = 60\%$$

Second, Player 2 reviews the percentage given by Player 1 and decides what percentage of Player 1's endowment to destroy. Suppose Player 2 chooses to destroy 50%.

$$\boxed{\text{Fraction destroyed by Player 2}} = 50\%$$

The final amounts earned in Experiments 2 and 3 are as follows:

$$\begin{aligned}
\text{Player 1's earnings} &= ¥200 \\
&= 1000 \times (100\% - 50\%) \quad - 1000 \times (100\% - 50\%) \times 60\% \\
\text{Player 2's earnings} &= ¥1{,}300 \\
&= 1000 \qquad\qquad\qquad\qquad + 1000 \times (100\% - 50\%) \times 60\%
\end{aligned}$$

Case 4 [for PTG+G and PTG+G+CLR]

- Player 1 chooses to give the endowment that Player 1 earned in Experiment 2.
- Player 2 chooses to destroy the endowment that Player 2 earned in Experiment 2.

In this case:

- Player 1's earnings are equal to their endowment minus the amount they gave to Player 2.
- Player 2's earnings are equal to the portion of their endowment that they did not destroy, plus the amount Player 1 gave to them.

- Specifically, the earnings are calculated using the following formulas:

$$\begin{aligned}
&\boxed{\text{Player 1's earnings}} \\
&= \boxed{\begin{array}{c}\text{Player 1's earning}\\ \text{in Experiment 2}\end{array}} \\
&\qquad - \boxed{\begin{array}{c}\text{Player 1's earning}\\ \text{in Experiment 2}\end{array}} \times \boxed{\begin{array}{c}\text{Fraction given}\\ \text{by Player 1}\end{array}}
\end{aligned}$$

$$\begin{aligned}
&\boxed{\text{Player 2's earnings}} \\
&= \boxed{\begin{array}{c}\text{Player 2's earning}\\ \text{in Experiment 2}\end{array}} \times \left(100\% - \boxed{\begin{array}{c}\text{Fraction destroyed}\\ \text{by Player 2}\end{array}}\right) \\
&\qquad + \boxed{\begin{array}{c}\text{Player 1's earning}\\ \text{in Experiment 2}\end{array}} \times \boxed{\begin{array}{c}\text{Fraction given}\\ \text{by Player 1}\end{array}}
\end{aligned}$$

Example. Assume that the earnings in Experiment 2 are as follows:

- Player 1's earnings in Experiment 2: ¥1,000
- Player 2's earnings in Experiment 2: ¥1,000

First, Player 1 decides what percentage of Player 1's endowment to give. Suppose Player 1 decides to give 60%.

$$\boxed{\begin{array}{c}\text{Fraction given}\\ \text{by Player 1}\end{array}} = 60\%$$

Second, Player 2 reviews the percentage given by Player 1 and decides what percentage of Player 2's endowment to destroy. Suppose Player 2 chooses to destroy 50%.

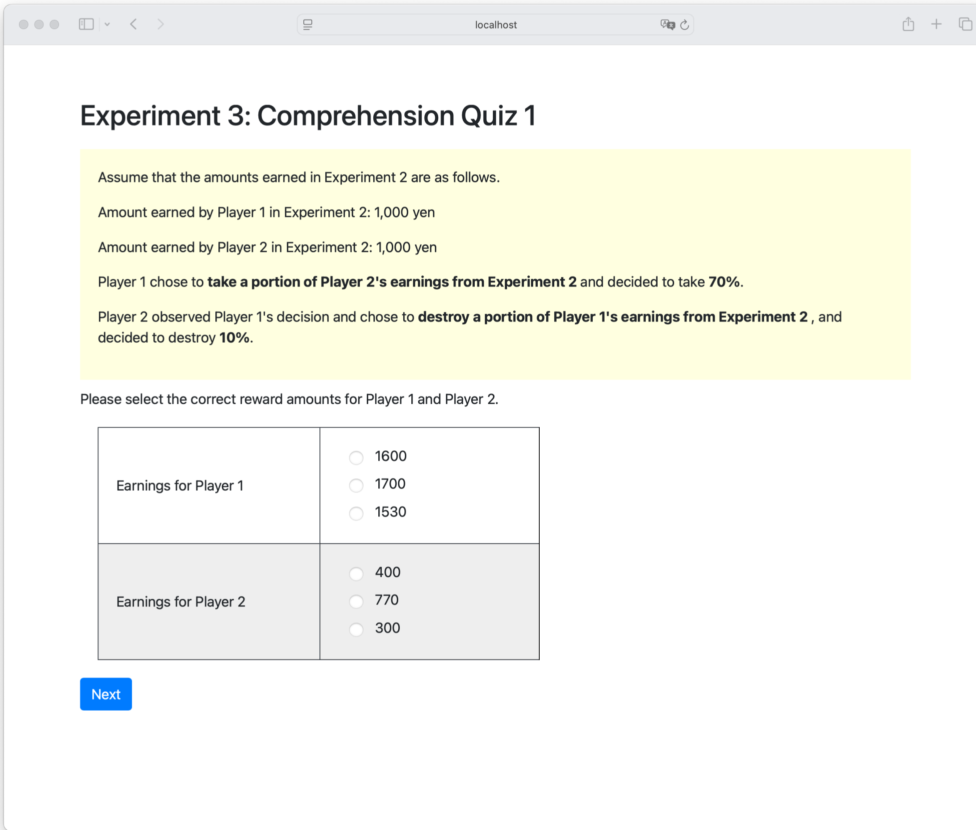
$$\boxed{\begin{array}{c}\text{Fraction destroyed}\\ \text{by Player 2}\end{array}} = 50\%$$

The final amounts earned in Experiments 2 and 3 are as follows:

$$\begin{aligned}
\text{Player 1's earnings} &= ¥400 \\
&= 1000 \qquad\qquad\qquad\qquad - 1000 \times 60\% \\
\text{Player 2's earnings} &= ¥1{,}100 \\
&= 1000 \times (100\% - 50\%) \quad + 1000 \times 60\%
\end{aligned}$$

Screenshots of the comprehension quiz

[For PTG+CLR and PTG+G+CLR]



localhost

Experiment 3: Comprehension Quiz 1

Assume that the amounts earned in Experiment 2 are as follows.

Amount earned by Player 1 in Experiment 2: 1,000 yen

Amount earned by Player 2 in Experiment 2: 1,000 yen

Player 1 chose to **take a portion of Player 2's earnings from Experiment 2** and decided to take **70%**.

Player 2 observed Player 1's decision and chose to **destroy a portion of Player 1's earnings from Experiment 2** , and decided to destroy **10%**.

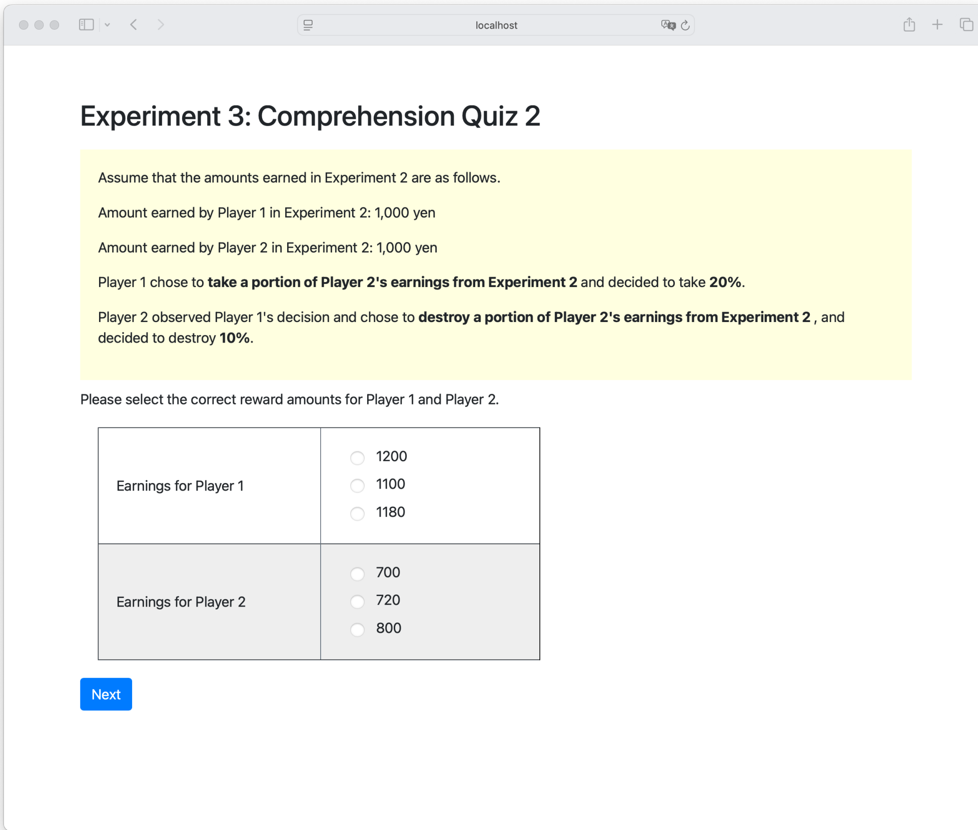
Please select the correct reward amounts for Player 1 and Player 2.

Earnings for Player 1	1600 1700 1530
Earnings for Player 2	400 770 300

Next

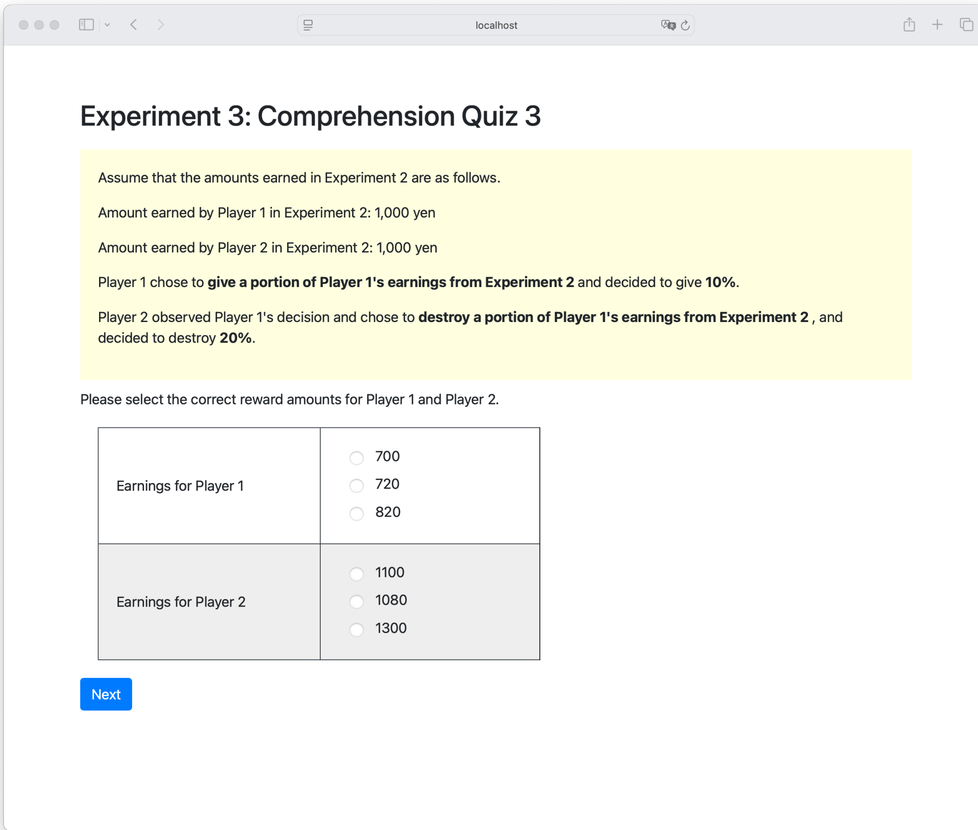
[Correct answers to the comprehension quiz 1: player 1's earnings are 1600, and player 2's earnings are 300.]

[For all treatments]



[Correct answers to the comprehension quiz 2: player 1's earnings are 1180, and player 2's earnings are 720.]

[For PTG+G+CLR]



localhost

Experiment 3: Comprehension Quiz 3

Assume that the amounts earned in Experiment 2 are as follows.

Amount earned by Player 1 in Experiment 2: 1,000 yen

Amount earned by Player 2 in Experiment 2: 1,000 yen

Player 1 chose to **give a portion of Player 1's earnings from Experiment 2** and decided to give **10%**.

Player 2 observed Player 1's decision and chose to **destroy a portion of Player 1's earnings from Experiment 2** , and decided to destroy **20%**.

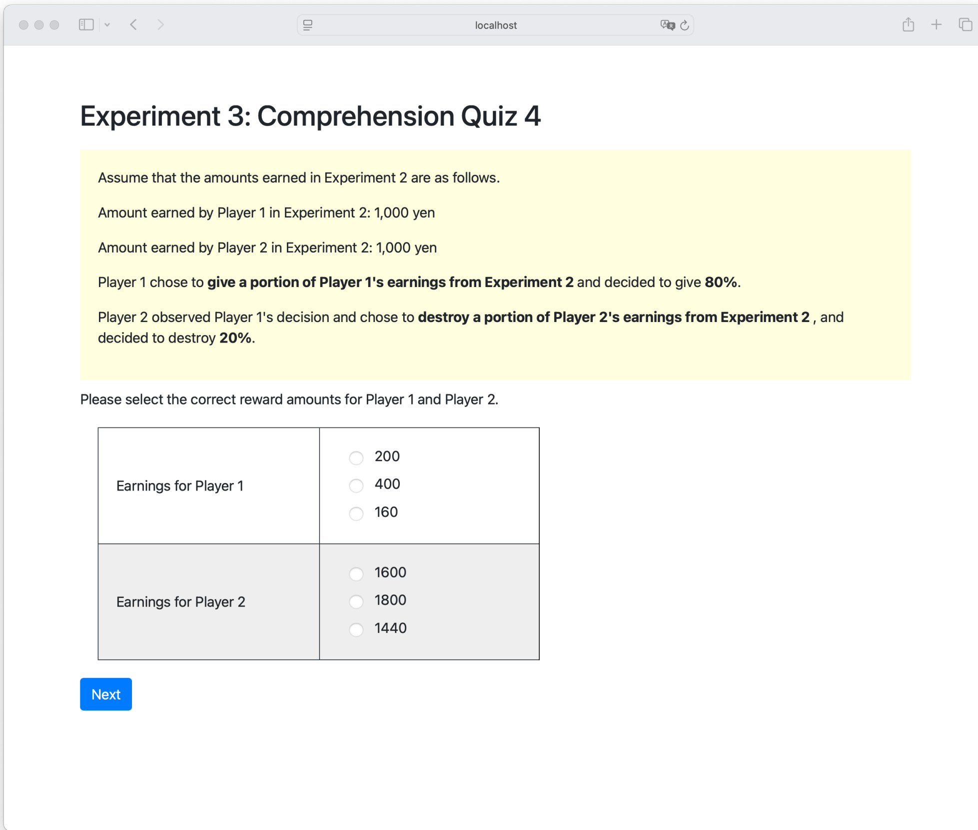
Please select the correct reward amounts for Player 1 and Player 2.

Earnings for Player 1	○ 700 ○ 720 ○ 820
Earnings for Player 2	○ 1100 ○ 1080 ○ 1300

Next

[Correct answers to the comprehension quiz 3: player 1's earnings are 720, and player 2's earnings are 1080.]

[For PTG+G and PTG+G+CLR]


localhost

Experiment 3: Comprehension Quiz 4

Assume that the amounts earned in Experiment 2 are as follows.

Amount earned by Player 1 in Experiment 2: 1,000 yen

Amount earned by Player 2 in Experiment 2: 1,000 yen

Player 1 chose to **give a portion of Player 1's earnings from Experiment 2** and decided to give **80%**.

Player 2 observed Player 1's decision and chose to **destroy a portion of Player 2's earnings from Experiment 2** , and decided to destroy **20%**.

Please select the correct reward amounts for Player 1 and Player 2.

Earnings for Player 1	200 400 160
Earnings for Player 2	1600 1800 1440

Next

[Correct answers to the comprehension quiz 4: player 1's earnings are 200, and player 2's earnings are 1600.]

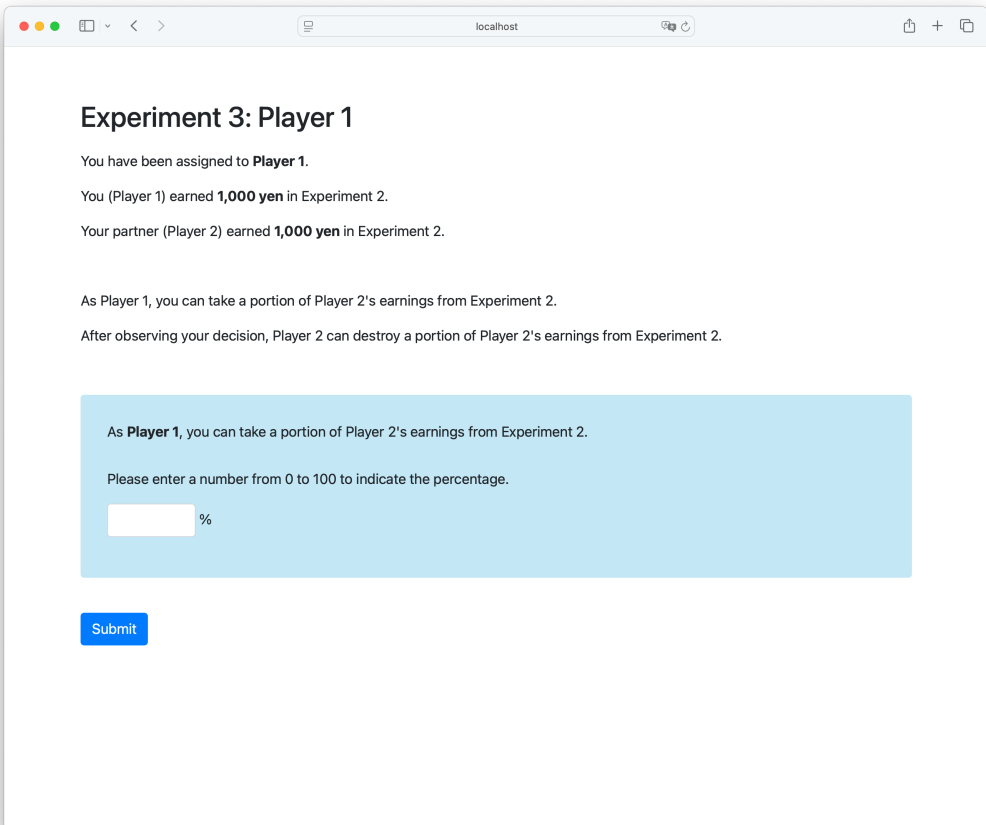
II.6 Survey on emotional state after the comprehension quiz for Experiment 3

[See Online Appendix II.2]

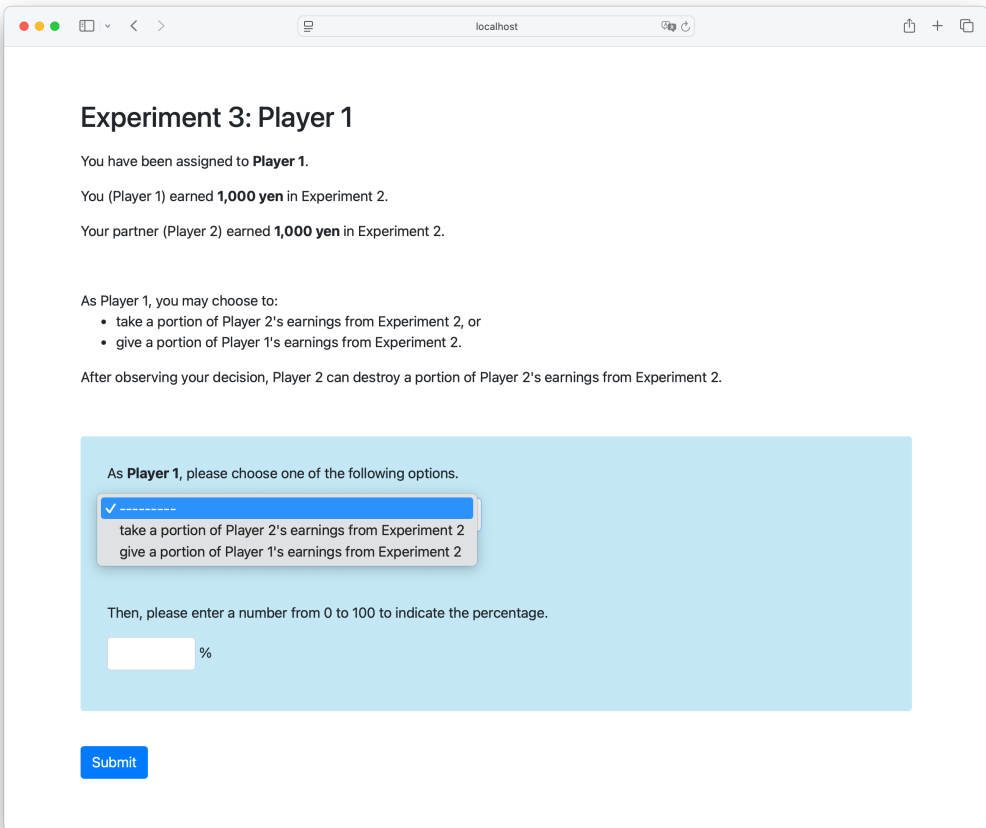
II.7 Main session of Experiment 3 (The power-to-take game)

Screenshots of the decision screen for proposers

[For PTG]



[For PTG+G]



localhost

Experiment 3: Player 1

You have been assigned to **Player 1**.

You (Player 1) earned **1,000 yen** in Experiment 2.

Your partner (Player 2) earned **1,000 yen** in Experiment 2.

As Player 1, you may choose to:

- take a portion of Player 2's earnings from Experiment 2, or
- give a portion of Player 1's earnings from Experiment 2.

After observing your decision, Player 2 can destroy a portion of Player 2's earnings from Experiment 2.

As **Player 1**, please choose one of the following options.

✓ ---------

take a portion of Player 2's earnings from Experiment 2

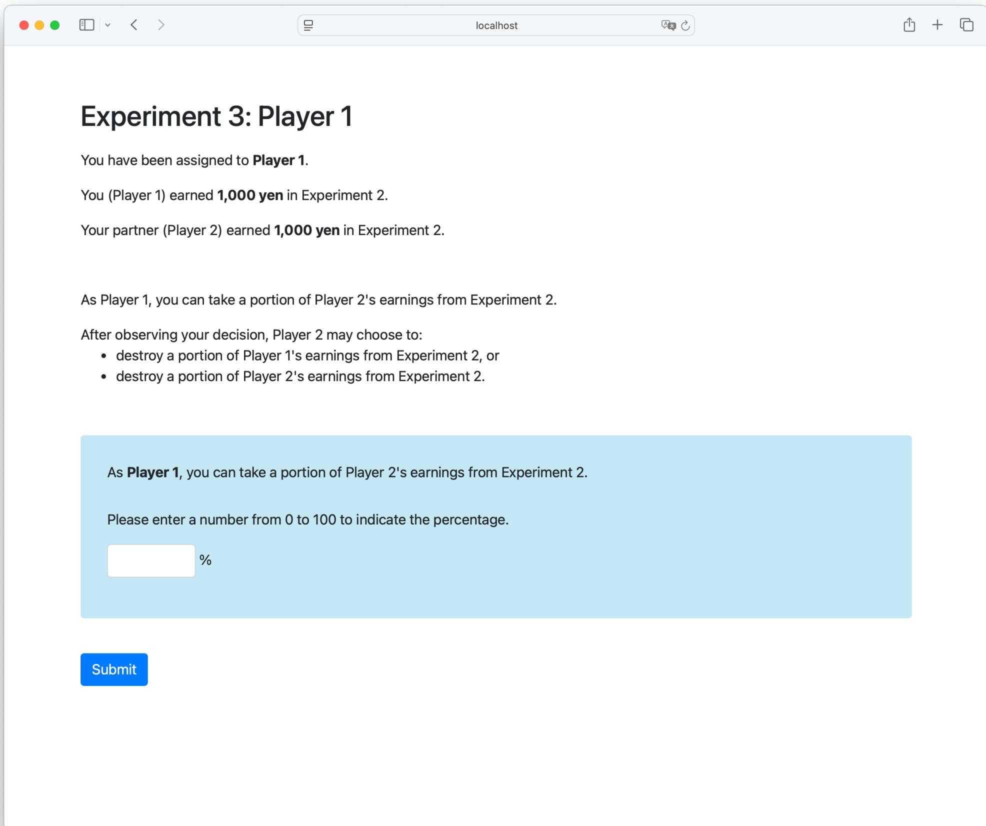
give a portion of Player 1's earnings from Experiment 2

Then, please enter a number from 0 to 100 to indicate the percentage.

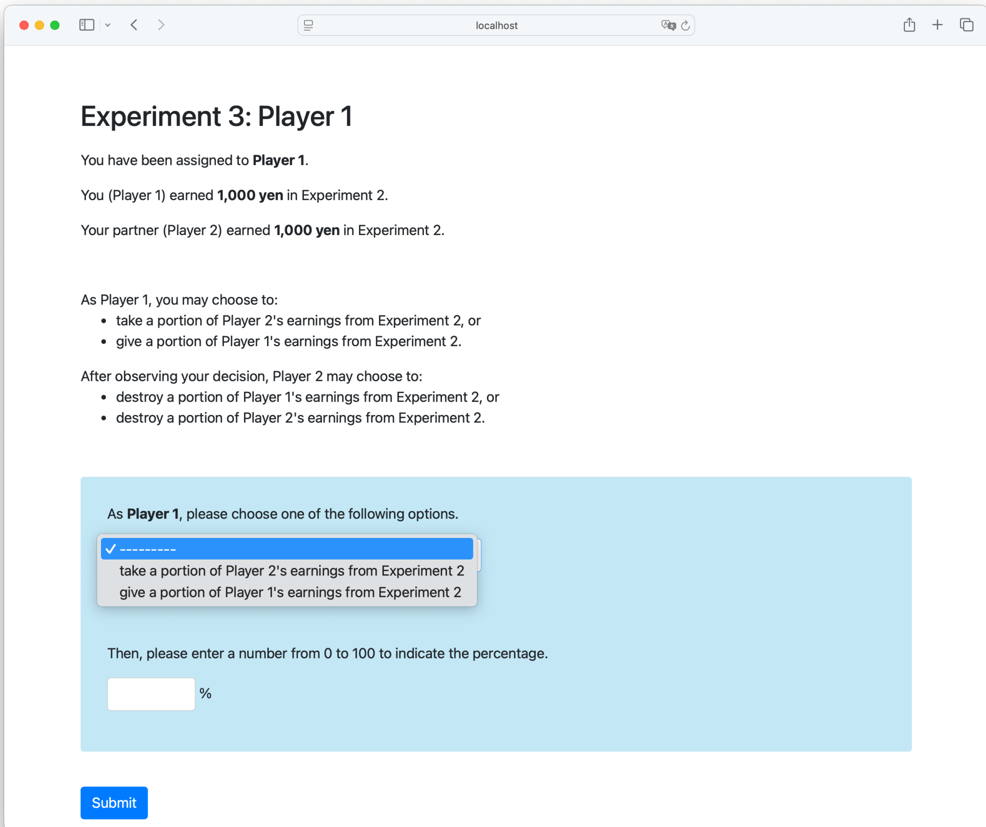
%

Submit

[For PTG+CLR]

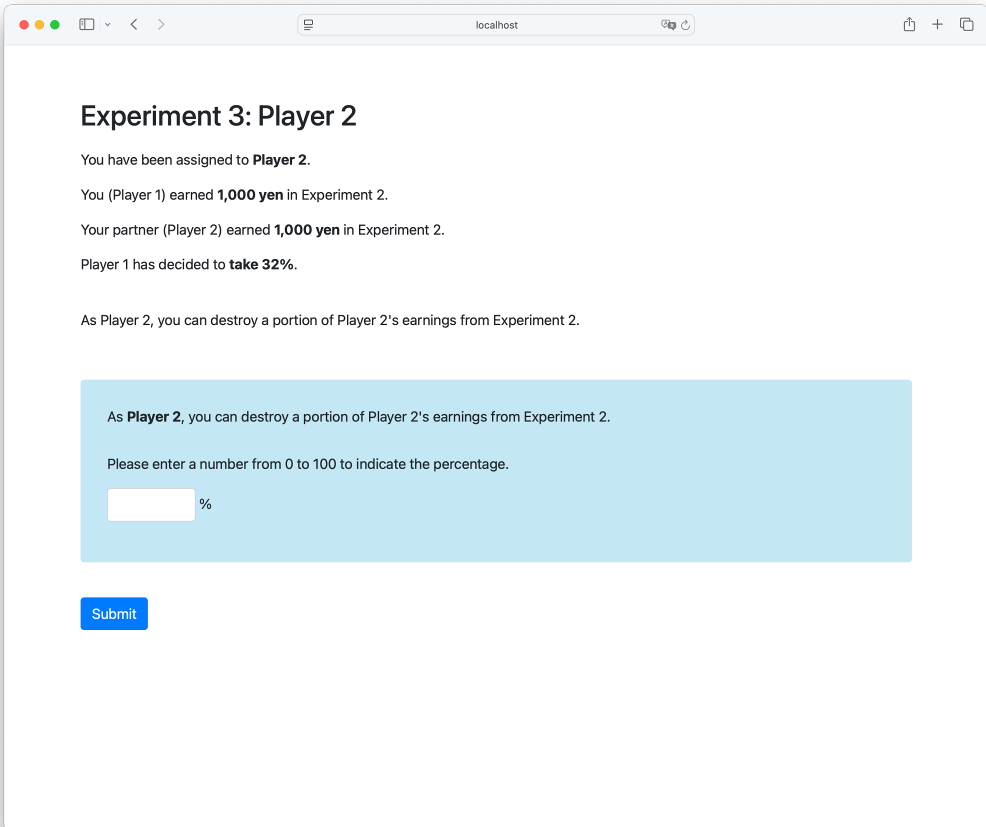

localhost
Experiment 3: Player 1
You have been assigned to Player 1.
You (Player 1) earned 1,000 yen in Experiment 2.
Your partner (Player 2) earned 1,000 yen in Experiment 2.
As Player 1, you can take a portion of Player 2's earnings from Experiment 2.
After observing your decision, Player 2 may choose to:
• destroy a portion of Player 1's earnings from Experiment 2, or
• destroy a portion of Player 2's earnings from Experiment 2.
As Player 1, you can take a portion of Player 2's earnings from Experiment 2.
Please enter a number from 0 to 100 to indicate the percentage.
%
Submit

[For PTG+G+CLR]

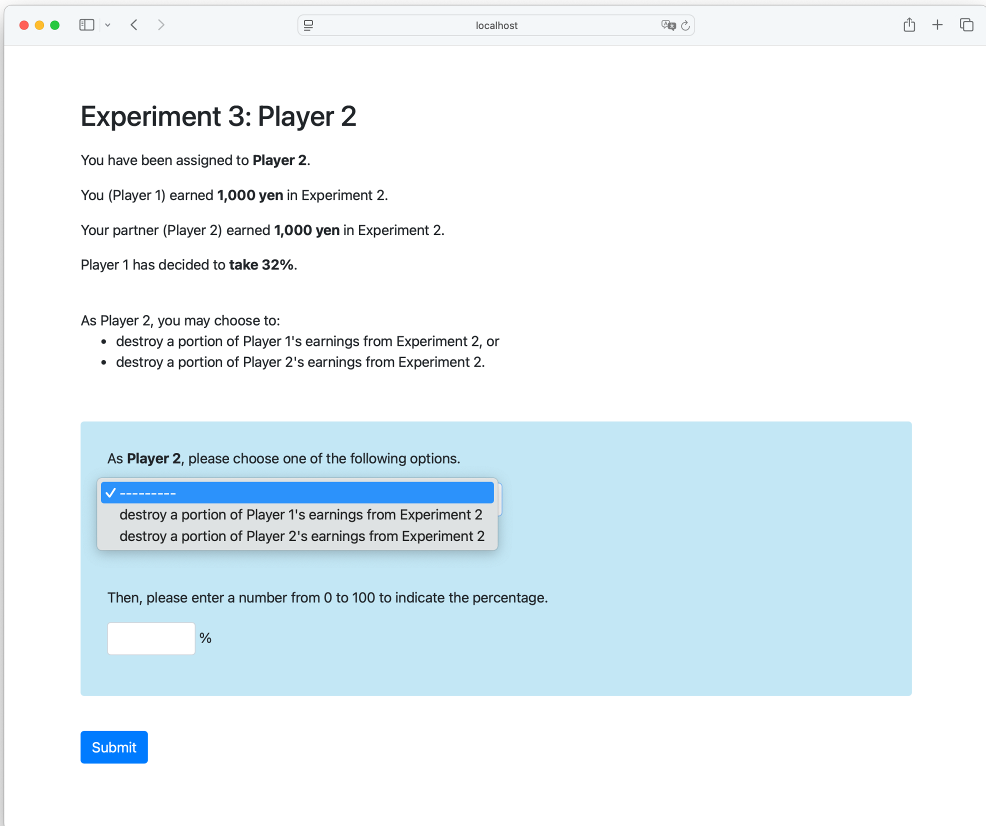

localhost
Experiment 3: Player 1
You have been assigned to Player 1.
You (Player 1) earned 1,000 yen in Experiment 2.
Your partner (Player 2) earned 1,000 yen in Experiment 2.
As Player 1, you may choose to:
• take a portion of Player 2's earnings from Experiment 2, or
• give a portion of Player 1's earnings from Experiment 2.
After observing your decision, Player 2 may choose to:
• destroy a portion of Player 1's earnings from Experiment 2, or
• destroy a portion of Player 2's earnings from Experiment 2.
As Player 1, please choose one of the following options.
----
take a portion of Player 2's earnings from Experiment 2
give a portion of Player 1's earnings from Experiment 2
Then, please enter a number from 0 to 100 to indicate the percentage.
%
Submit

Screenshots of the decision screen for responders

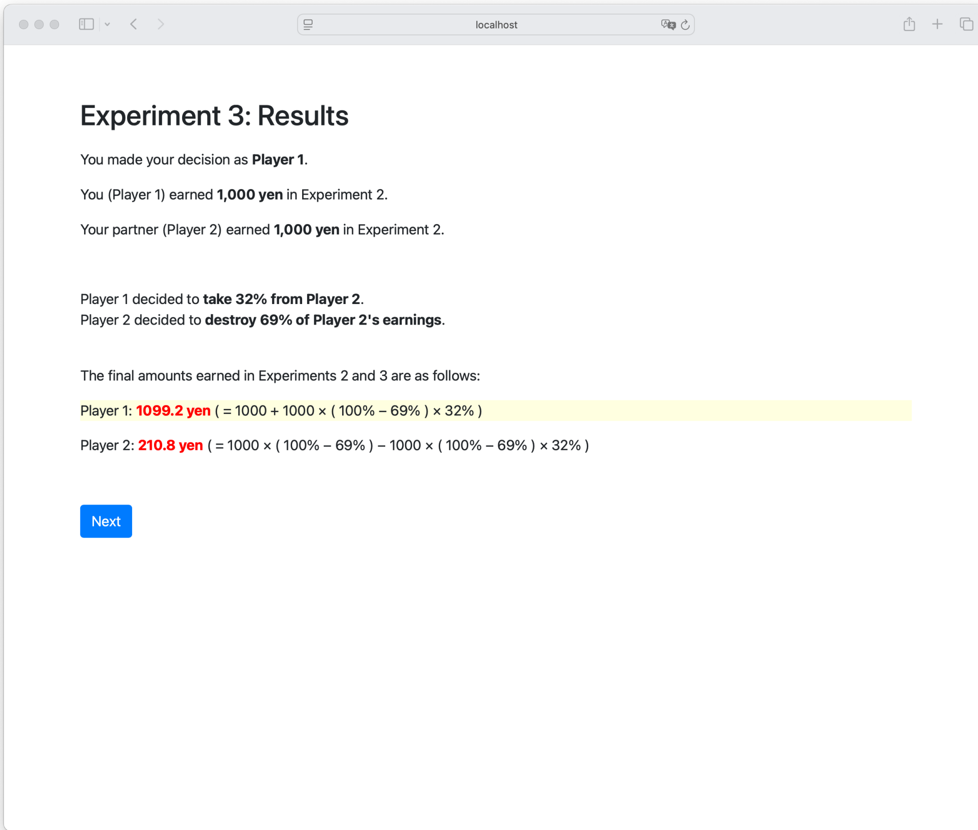
[For PTG and PTG+G]



[For PTG+CLR and PTG+G+CLR]



Screenshot of the feedback

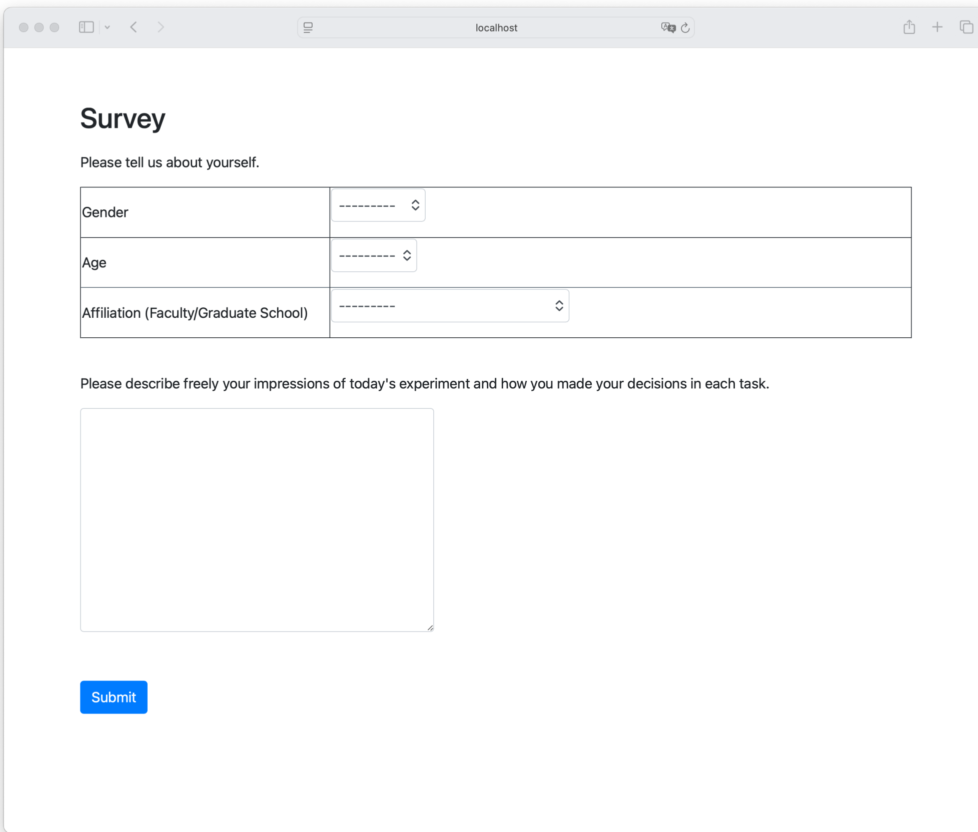


II.8 Surveys after the feedback on the results of Experiment 3

Survey on emotional state

[See Online Appendix II.2]

Survey on participant characteristics



localhost

Survey

Please tell us about yourself.

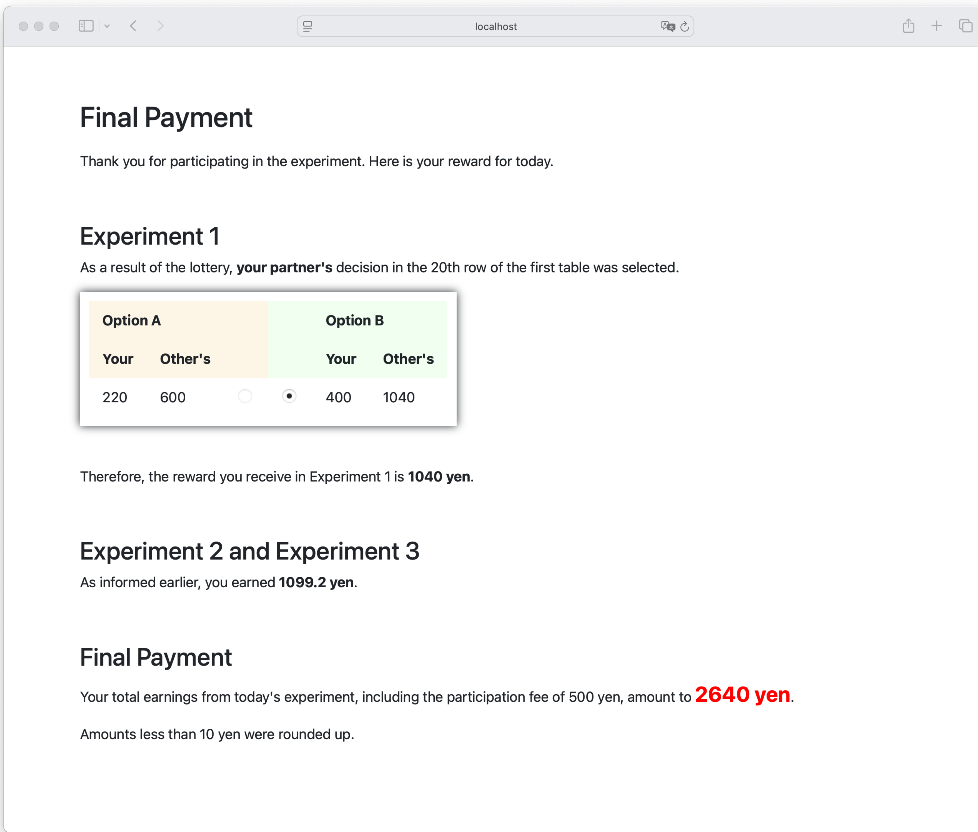
Gender	---------
Age	---------
Affiliation (Faculty/Graduate School)	---------

Please describe freely your impressions of today's experiment and how you made your decisions in each task.

Submit

II.9 Feedback on the final payment

Screenshot



Final Payment

Thank you for participating in the experiment. Here is your reward for today.

Experiment 1

As a result of the lottery, **your partner's** decision in the 20th row of the first table was selected.

Option A				**Option B**	
Your	**Other's**			**Your**	**Other's**
220	600	○	◉	400	1040

Therefore, the reward you receive in Experiment 1 is **1040 yen**.

Experiment 2 and Experiment 3

As informed earlier, you earned **1099.2 yen**.

Final Payment

Your total earnings from today's experiment, including the participation fee of 500 yen, amount to **2640 yen**.

Amounts less than 10 yen were rounded up.

III Balance check

In this appendix, we demonstrate that our samples are balanced across the four treatments with respect to gender composition, measured degrees of inequality aversion (α and β), and initial emotional state. We also report the results of the real effort task.

III.1 Gender composition

Table III.1 shows the gender composition across the four treatments. The composition does not differ significantly among treatments (Pearson's $\chi^2(3) = 1.2447$, $p = 0.742$).

III.2 Inequity aversion

Table III.2 reports the mean (standard deviation) of the measured parameters α and β for each treatment. There is no statistically significant difference in the degree of inequality aversion across the four treatments. Note that four participants (two in each of the PTG+G and PTG+CLR treatments) were excluded from the β estimation because they selected the boundary of the list as their switching point.

Figure III.1 shows the joint distribution of the parameters α and β as a bubble chart. The size of each bubble corresponds to the number of observations at its center. For most participants, (α, β) take one of two values: $(0, 0.111)$—with 19, 14, 11, and, 14 observations in PTG, PTG+G,

Table III.1: Gender composition by treatment

	Total		Proposer		Responder	
	Male	Female	Male	Female	Male	Female
PTG	62	23	31	12	31	11
PTG+CLR	55	26	27	13	28	13
PTG+G	55	25	25	15	30	10
PTG+G+CLR	52	28	25	15	27	13
p-value (χ^2 test)	0.742		0.756		0.836	

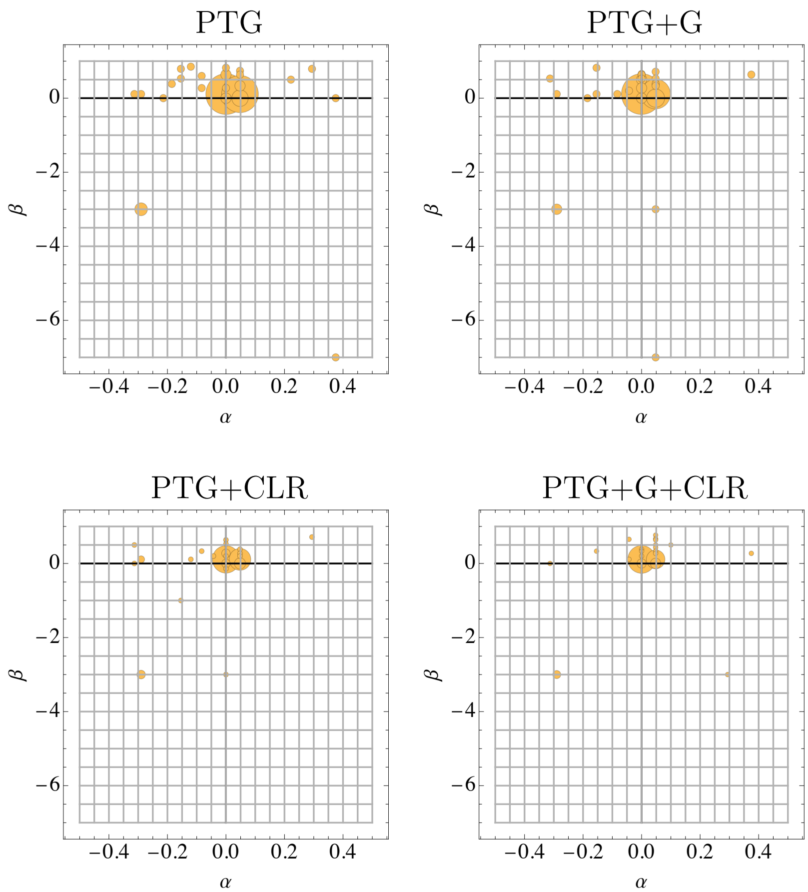
Note: Two participants—one in the responder role under the PTG treatment and one in the proposer role under the PTG+CLR treatment—did not specify their gender binary and are therefore excluded from this table.

Table III.2: Average measured values of α and β by treatment

	α	β
PTG	0.005	−0.006
	(0.113)	(0.985)
PTG+CLR	−0.011	−0.032
	(0.101)	(0.709)
PTG+G	−0.001	−0.040
	(0.092)	(1.030)
PTG+G+CLR	0.009	0.001
	(0.092)	(0.710)
p-value (KW test)	0.631	0.878

Note: Standard deviations are shown in parentheses.

Figure III.1: Distributions of elicited (α, β)



PTG+CLR, and PTG+G+CLR, respectively—or $(0.048, 0.111)$—with 13, 8, 13, and 9 in PTG, PTG+G, PTG+CLR, and PTG+G+CLR, respectively.

III.3 Real effort task

As noted in Section 2, we set the reward for the real effort task so that every participant could receive a maximum endowment of 1,000 JPY for use in the PTG. Table III.3 summarises the outcomes of the real effort task in terms of the number of correct answers.

As one can observe, although participants were required to answer 40 questions correctly to obtain the full endowment, they answered, on average, more than 100 questions correctly. The results of a two-way ANOVA indicate that there was a significant difference between treatments ($p = 0.008$), but no significant difference across roles ($p = 0.685$). Recall that the real effort task was completed before roles in the PTGs were assigned.

Table III.4 examines players' decision, controlling for the number of correct answers in the real effort task. The estimated coefficients on effort are not significantly different from zero, and the treatment dummy coefficients closely match those reported in Tables 4 and 5.

Table III.3: Average outcomes of the real effort task

	Proposer	Responder
PTG	85.33	83.02
	(22.71)	(20.27)
PTG+CLR	89.63	92.1
	(20.37)	(23.07)
PTG+G	97.1	95.48
	(24.05)	(23.46)
PTG+G+CLR	94.38	91.62
	(26.57)	(26.59)
p-value (KW test)	0.116	0.094

Note: Standard deviations are shown in parentheses.

Table III.4: Effect of real effort task outcome

	Proposer's decision		Responder's decision			
	Increase in proposer's payoff due to proposer $tY_2 - gY_1$		Decrease in proposer's payoff due to responder $(1-g)rY_1 + tdY_2$		Decrease in responder's payoff due to responder $grY_1 + (1-t)dY_2$	
	(1)	(2)	(3)	(4)	(5)	(6)
effort	−0.127	0.506	−0.093	−1.364	−0.047	−0.073
	(1.264)	(1.263)	(1.494)	(1.030)	(0.120)	(0.122)
$I_{+\text{CLR}}$		52.568		527.499***		−4.401
		(82.363)		(67.234)		(7.985)
$I_{+\text{G}}$		63.708		99.410		13.660*
		(83.575)		(68.260)		(8.107)
$I_{+\text{G+CLR}}$		−179.595**		454.242***		3.285
		(82.798)		(68.494)		(8.135)
$t - g$				7.011***		−0.039
				(0.647)		(0.077)
const.	537.030***	494.751***	345.007**	−175.326*	12.084	13.488
	(119.481)	(119.320)	(139.567)	(102.402)	(11.258)	(12.162)
N	164	164	164	164	164	164
adj. R^2	−0.006	0.042	−0.006	0.541	−0.005	0.005

Note: ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors are shown in parentheses.

IV Additional analyses

IV.1 Proposers' behaviour

We estimate a probit model to analyse proposers' propensity to take from responders—regardless of amount—under each treatment condition. The dependent variable is an indicator variable that equals one if the proposer chooses a strictly positive take rate ($t > 0$).

In Column (1) of Panel (a) in Table IV.1, we include only the treatment indicators. In Panel (b), we conduct Wald tests on the coefficients to compare treatments. These results confirm that, relative to the other three treatments, the PTG+G+CLR treatment significantly reduces the likelihood that a proposer will take.

In the second specification (Column (2)), we further control for each participant's degree of inequality aversion by including parameters α and β. Adding these controls improves model fit, as evidenced by increases in pseudo R^2 and reductions in the Akaike Information Criterion (AIC), but the treatment-indicator coefficients remain virtually unchanged.

Next, we analyse how the treatment conditions affect proposers' decisions among those who did not give to the responder, including those who chose neither to give nor to take). To this end, we employ a Tobit model, using the increase in the proposer's payoff, $tY_2 - 0 \cdot Y_1$, as the dependent variable; this variable is censored below at zero (no taking) and above at the responder's initial endowment (full taking).

Table IV.1: Probit and Tobit model analyses of proposers' decision

(a) Regression results

	Probit		Tobit excluding giving proposers	
	(Taking)		(Increase in proposer's payoff)	
	(1)	(2)	(3)	(4)
$I_{+\mathrm{CLR}}$	−0.371	−0.435	104.150	107.916
	(0.353)	(0.362)	(109.365)	(110.212)
$I_{+\mathrm{G}}$	−0.041	−0.060	99.880	81.092
	(0.379)	(0.388)	(108.808)	(109.856)
$I_{+\mathrm{G+CLR}}$	−1.133***	−1.209***	93.930	83.759
	(0.333)	(0.347)	(123.046)	(125.292)
α		1.211		132.703
		(1.385)		(462.566)
β		−0.554		−47.620
		(0.357)		(58.582)
const.	1.322***	1.413***	525.727***	528.530***
	(0.266)	(0.282)	(75.028)	(76.173)
N	164	162	152	150
pseudo R^2	0.107	0.133	0.001	0.001
AIC	152.648	151.537	1680.975	1666.791

Note: ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors are shown in parentheses.

(b) P-values for pairwise comparisons based on regression (1) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	0.293	0.914	< 0.001
PTG+CLR	-	0.353	0.013
PTG+G	-	-	0.001

(c) P-values for pairwise comparisons based on regression (3) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	0.342	0.360	0.446
PTG+CLR	-	0.970	0.935
PTG+G	-	-	0.962

Columns (3) and (4) of Panel (a) in Table IV.1 report two Tobit regressions. The first includes only the treatment dummies; the second also incorporates the degree of inequality aversion. In both regressions, no coefficients is statistically significant, the pseudo R^2 values are low, and overall model fit is poor. In Panel (c), we conduct Wald tests on the coefficients of Columns (3) to compare treatments—pairwise comparisons reveal no significant differences.

Taken together with the probit results, these findings suggest that the PTG+G+CLR treatment significantly increases the probability that proposers choose to 'take,' but does not significantly affect the amount taken by those who do so.

IV.2 Responders' behaviour

We estimate a probit model to test the effect of the treatments on the probability that a responder chooses to destroy. The dependent variable is an indicator equal to one if the responder's destroy rate is strictly positive ($d > 0$).

Columns (1)–(3) in Panel (a) of Table IV.2 report three probit regressions. Column (1), includes only the treatment indicators; Column (2) adds the proposer's $(t - g)$ as a control; and Column (3) further controls for the responder's degree of inequality aversion. Panels (b) and (c) of Table IV.2 report the p-values from Wald tests for pairwise treatment comparisons based on Columns (1) and (2), respectively. Across all three specifications, only the PTG+G treatment is statistically significant at the 5% level—that is, only under PTG+G are responders more likely to choose the destroy action.

Next, we examined the effect of the treatment conditions on the decisions of responders who did not choose retaliation, including those who neither retaliated nor destroyed any of the proposer's payoff. We used a Tobit model with two dependent variables: reductions in the proposer's payoff, $0 \cdot (1 - g)Y_1 + tdY_2$ (Columns (4)–(6)) and reductions in the responder's payoff, $0 \cdot gY_1 + (1 - t)dY_2$ (Columns (7)–(9)), both of which were censored from below at zero (no destruction) and from above at the maximum possible destruction—namely, tY_2 for Columns (4)–(6) and $(1 - t)Y_2$ for Columns (7)–(9). In Columns (4) and (7), only the treatment indicators are included; Columns (5) and (8) add the control for the proposer's $(t - g)$; and Columns (6) and (9) further controlled for the responder's degree of inequality aversion. Panels (d)–(g) of Table IV.2 reported the p-values from Wald tests for pairwise treatment comparisons based on Columns (4), (5), (7), and (8).

In Columns (4)–(6), significant effects at the 5% level were observed in both the PTG+CLR and PTG+G+CLR treatments. In Columns (7)–(9), a significant effect at the 1% level was found in the PTG+G+CLR treatment, but only when the proposer's was controlled.

IV.3 Estimation of a reciprocity-based model

To better understand the behaviour observed across the four treatments, we employ the 'tractable' framework of reciprocity and fairness proposed

Table IV.2: Probit and Tobit model analyses of responders' decision

(a) Regression results

	Probit			Tobit excluding retaliating responders					
	(Destroying)			(Decrease in proposer's payoff)			(Decrease in responder's payoff)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$I_{+\mathrm{CLR}}$	−0.574	−0.723	−0.690	2064.555**	2107.367**	2322.089**	101.401	204.087	179.678
	(0.409)	(0.440)	(0.448)	(951.879)	(999.152)	(1137.277)	(78.867)	(129.557)	(123.028)
$I_{+\mathrm{G}}$	0.764**	0.736**	0.713**	1119.745*	1214.793*	1212.094*	130.146**	138.062*	114.058*
	(0.312)	(0.327)	(0.341)	(594.476)	(648.239)	(724.909)	(59.036)	(70.927)	(68.683)
$I_{+\mathrm{G+CLR}}$	−0.199	0.030	0.016	2683.052**	2678.309**	2992.625**	129.522*	456.828***	455.686***
	(0.360)	(0.372)	(0.382)	(1074.942)	(1111.912)	(1269.899)	(71.105)	(160.222)	(159.830)
$t-g$		0.012***	0.011***		−4.948	−5.489		5.589***	5.474***
		(0.004)	(0.004)		(7.414)	(8.348)		(1.828)	(1.826)
α			−0.284			50.656			−286.972
			(1.301)			(2731.128)			(267.234)
β			−0.225*			−244.110			−29.318
			(0.129)			(219.846)			(18.459)
const.	−1.082***	−1.798***	−1.745***	−1364.649**	−1176.503*	−1326.328*	−187.511***	−573.256***	−555.413***
	(0.238)	(0.351)	(0.360)	(574.030)	(620.166)	(709.265)	(61.622)	(177.094)	(173.677)
N	164	164	161	112	112	110	112	112	110
pseudo R^2	0.116	0.185	0.197	0.055	0.056	0.064	0.028	0.175	0.179
AIC	137.668	129.563	128.251	315.104	316.592	300.523	271.834	234.236	234.473

Note: ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors are shown in parentheses.

Table IV.2: Probit and Tobit model analyses of responders' decision (cont'd)

(b) P-values for pairwise comparisons based on regression (1) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	0.160	0.014	0.580
PTG+CLR	-	< 0.001	0.381
PTG+G	-	-	0.004

(c) P-values for pairwise comparisons based on regression (2) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	0.101	0.024	0.936
PTG+CLR	-	< 0.001	0.105
PTG+G	-	-	0.042

(d) P-values for pairwise comparisons based on regression (4) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	0.032	0.062	0.014
PTG+CLR	-	0.237	0.508
PTG+G	-	-	0.076

(e) P-values for pairwise comparisons based on regression (5) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	0.037	0.064	0.018
PTG+CLR	-	0.287	0.565
PTG+G	-	-	0.112

(f) P-values for pairwise comparisons based on regression (7) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	0.201	0.030	0.071
PTG+CLR	-	0.685	0.731
PTG+G	-	-	0.992

(g) P-values for pairwise comparisons based on regression (8) (Wald test)

	PTG+CLR	PTG+G	PTG+G+CLR
PTG	0.118	0.054	0.005
PTG+CLR	-	0.571	0.096
PTG+G	-	-	0.021

by Cox et al. (2007) to estimate the relative weight participants assign to their opponent's payoff when choosing their actions.[2] We use this outcome-based reciprocity model because it provides a quantitative representation of how players value others' payoffs. We posit that the weight placed on the opponent's payoff is driven by participants' emotional states in each treatment and is, therefore, influenced both by the set of actions available to the two players and by the proposer's chosen action when interacting with the responder.

IV.3.1 Framework

Assume that a player's social preferences are represented by the following constant elasticity of substitution (CES) utility function:

$$u(\pi_{self}, \pi_{opp}) = \begin{cases} \pi_{self}^{\rho} + \theta\, \pi_{opp}^{\rho} & \text{for} \quad \rho \in (0, 1], \\ \log(\pi_{self}) + \theta\, \log(\pi_{opp}) & \text{for} \quad \rho = 0, \end{cases}$$

where π_{self} and π_{opp} denote the payoffs to the player and the opponent, respectively. Specifically, for proposers (player 1) we set $self = 1$, $opp = 2$; for responders (player 2) we set $self = 2$, $opp = 1$.

The parameter θ reflects the player's 'emotional status' (Cox et al., 2007),

[2] Online Appendix V presents analyses of emotional changes and their relationship to behaviour. We did not find any significant correlation between these emotional shifts and the responder's behaviour.

which may be influenced both by the players' available action sets and by the opponent's prior choices. The parameter ρ governs the curvature of the indifference curves. As Thöni (2015) and Inukai et al. (2024) show, when $\rho < 0$—that is, when the elasticity of substitution falls below that of a Cobb–Douglas utility—the weighting parameter θ becomes uninterpretable. We therefore exclude all cases with $\rho < 0$.

Under this representation, the marginal rate of substitution between a player's own payoff and the opponent's payoff is given by $\partial\pi_{opp}/\partial\pi_{self} = \theta^{-1}(\pi_{opp}/\pi_{self})^{1-\rho}$. Hence, when $\pi_{self} = \pi_{opp}$, $\theta = \partial\pi_{self}/\partial\pi_{opp}$, which captures the rate at which a player trades off their own payoff for the opponent's payoff when both payoffs are equal. We now apply this framework to representative players—aggregated across all participants by role (proposers vs. responders) and by treatment—and estimate their preferences.

IV.3.2 Estimating aggregated proposers' preferences

We estimate proposers' preferences based on the interim payoffs prior to responders' actions. We interpret the parameter θ and the observed proposer choices in light of proposers' likely expectations about responder retaliation. These expectations shape the proposer's decision-making process, even though they are not explicitly included in the utility function used for estimation.

For the proposer (player 1), the payoffs to players 1 and 2—resulting from the proposer's choice but before the responder's decision—are

$$\pi_1 = (1-g)Y_1 + tY_2, \quad \pi_2 = gY_1 + (1-t)Y_2,$$

with initial endowments $Y_1 = Y_2 = 1{,}000$. We focus on the proposer (player 1), so we replace π_{self} and π_{opp} with π_1 and π_2, respectively. Here, t denotes the proportion the proposer takes from the responder, and g denotes the proportion the responder gives to the proposer. These two actions cannot occur simultaneously.

When the proposer does not have the option to give (i.e., $g = 0$), they choose t from the set of integer percentages from 0 to 100, inclusive, yielding 101 possible actions. Under a multinomial logit specification, the probability that the proposer selects $t = x^t$ is

$$\Pr(t = x^t) = \frac{\exp\left[u\big(\pi_1(t = x^t, g = 0), \pi_2(t = x^t, g = 0)\big)\right]}{\sum_{k=0}^{100} \exp\left[u\big(\pi_1(t = k/100, g = 0), \pi_2(t = k/100, g = 0)\big)\right]}.$$

When the proposer has the option to give, they choose either t or g—each an integer percentage from 0 to 100, subject to the constraint that t and g cannot both be strictly positive—resulting in 201 possible cases. For a give choice $(t, g) = (x^t, x^g)$, the numerator of the choice probability is $\exp[u(\pi_1(t = x^t, g = 0), \pi_2(t = x^t, g = 0))]$ for $t \geq 0$, and $\exp[u(\pi_1(t = 0, g =$

$x^g), \pi_2(t = 0, g = x^g))]$ for $g \geq 0$, while the denominator is replaced by

$$\begin{aligned}
&\exp\Big[u\big(\pi_1(t=0, g=0), \pi_2(t=0, g=0)\big)\Big] \\
&\quad + \sum_{k=1}^{100} \exp\Big[u\big(\pi_1(t=k/100, g=0), \pi_2(t=k/100, g=0)\big)\Big] \\
&\quad + \sum_{k=1}^{100} \exp\Big[u\big(\pi_1(t=0, g=k/100), \pi_2(t=0, g=k/100)\big)\Big]
\end{aligned}$$

in both cases.

Let $x_1, x_2, \ldots, x_N$ denote the individual decision data from the N proposers, the log-likelihood function is

$$\log L = \sum_{i=1}^{N} \log \Pr(x_i).$$

We estimate the weighting parameter θ for each treatment, holding the curvature parameter ρ fixed at eight distinct values so that the logarithm of the elasticity of substitution, $-\log(1-\rho)$, takes the values ({0.0, 0.5, 1.0, 1.5, 2.0, 3.0, 4.0, ∞}).

Table IV.3 presents the estimated values of θ for these eight values of ρ across four treatments. The log-likelihood is maximised at $\rho = 0.0$ among the values considered; hence, we focus on the results at this value.

The estimated values of θ are negative in all four treatments and are statistically significant in the three treatments other than PTG. There is no significant difference in the estimated values of θ among these three treat-

Table IV.3: Estimation results for the proposers' parameter θ

$-\log(1-\rho)$	ρ	PTG	PTG+CLR	PTG+G	PTG+G+CLR	$-\log L$
0.0	0.000	–0.063	–0.299***	–0.271***	–0.297***	
		(0.14)	(0.04)	(0.04)	(0.05)	
		197.96	*168.32*	*181.17*	*184.77*	*732.22*
0.5	0.393	–0.237	–2.450**	–3.002***	–2.007***	
		(0.71)	(0.92)	(0.51)	(0.69)	
		198.08	*180.85*	*185.56*	*198.99*	*763.48*
1.0	0.632	0.115	–1.266	–2.172***	–0.904*	
		(0.58)	(0.87)	(0.45)	(0.47)	
		198.11	*185.48*	*186.68*	*202.67*	*772.94*
1.5	0.777	0.290	–0.765	–1.831***	–0.536	
		(0.58)	(0.76)	(0.42)	(0.39)	
		198.11	*186.60*	*186.95*	*203.48*	*775.13*
2.0	0.865	0.396	–0.501	–1.655***	–0.370	
		(0.55)	(0.77)	(0.40)	(0.35)	
		198.11	*186.93*	*187.03*	*203.76*	*775.82*
3.0	0.950	0.502	–0.262	–1.500***	–0.236	
		(0.52)	(0.79)	(0.40)	(0.31)	
		198.11	*187.06*	*187.08*	*203.95*	*776.20*
4.0	0.982	0.543	–0.177	–1.446***	–0.193	
		(0.51)	(0.74)	(0.40)	(0.29)	
		198.11	*187.07*	*187.09*	*204.01*	*776.28*
∞	1.000	0.566	–0.128	–1.415***	–0.169	
		(0.52)	(0.76)	(0.37)	(0.29)	
		198.11	*187.07*	*187.10*	*204.04*	*776.31*

Note: We fixed the curvature parameter ρ at eight distinct values and estimated the weighting parameter θ for each treatment. In each cell, the top row reports the estimated value of θ, the middle row its standard error, and the bottom row the negative log-likelihood. Standard errors were calculated as the standard deviation of the estimates across bootstrapped samples. The first column of the table reports the logarithm of the elasticity of substitution. When $\rho = 0$ (i.e., the Cobb–Douglas utility function), the total negative log-likelihood is minimised, indicating the best fit. ***, ** ,*: statistically significantly different from zero at 1%, 5%, and 10% significance level.

ments. On average, proposers are willing to reduce the responder's payoff by almost 3 points to increase their own payoff by 10 points.

In PTG+G+CLR, the presence of costless retaliation might create a stronger expectation—or social norm—for proposers to give more or take less, influencing their behaviour despite a similar underlying reciprocal preference. This suggests that the opportunity to give does not necessarily translate into more generous behaviour unless it is coupled with a clear expectation or likelihood of retaliation. The introduction of costless retaliation adds a new dynamic whereby responders can retaliate at no cost. Consequently, proposers may give more or to take less to avoid potential retaliation, even if their overall reciprocal preferences remain unchanged—an indication of an implicit social norm activated by the costless-retaliation treatment.

IV.3.3 Estimating aggregated responders' preferences

We now turn to the responder. To focus on the responder (player 2), we have replaced the labels π_{self} and π_{opp} in Subsection IV.3.1 with π_2 and π_1, respectively. The payoffs for the responder (player 2) and the proposer (player 1), after the responder's decision, are given by

$$\pi_2 = (1-r)gY_1 + (1-t)(1-d)Y_2,$$

$$\pi_1 = (1-r)(1-g)Y_1 + t(1-d)Y_2.$$

In these expressions, d denotes the proportion of the responder's own payoff that is destroyed, and r denotes the proportion of the proposer's payoff that is destroyed in retaliation. Importantly, d and r cannot both be positive simultaneously. Moreover, at the time the responder makes their decision, t and g are already known. Therefore, we hypothesise that θ depends on the proposer's observed choice. For simplicity, we set

$$\theta = b\,(t - g),$$

where the coefficient b may vary by treatment.[3]

When the responder cannot retaliate (i.e., $r = 0$), they choose d from the set of integer percentages from 0 to 100, inclusive, yielding 101 possible actions. Under a multinomial logit specification, the probability that the responder selects $d = y^d$ is

$$\Pr(d = y^d \mid x) = \frac{\exp\left[u\big(\pi_2(d = y^d, r = 0),\, \pi_1(d = y^d, r = 0) \mid x\big)\right]}{\sum_{k=0}^{100} \exp\left[u\big(\pi_2(d = k/100, r = 0),\, \pi_1(d = k/100, r = 0) \mid x\big)\right]}$$

where x denotes the proposer's decision.

When the responder can retaliate, they choose either d or r—each an integer percentage between 0 and 100, subject to the constraint that d and r cannot both be strictly positive—resulting in 201 possible outcomes. The

[3]We also considered the specification $\theta = a + b\,(t - g)$, but the estimation procedure failed to converge.

numerator in the choice probability for $(d, r) = (y^d, y^r)$ is $\exp[u(\pi_2(d = y^d, r = 0), \pi_1(d = y^d, r = 0) \mid x)]$ for $d \geq 0$, and $\exp[u(\pi_2(d = 0, r = y^r), \pi_1(d = 0, r = y^r) \mid x)]$ for $r \geq 0$, while the denominator is replaced by

$$\begin{aligned}\exp\Big[u\big(&\pi_2(d=0,r=0),\pi_1(d=0,r=0)\mid x\big)\Big]\\ &+\sum_{k=1}^{100}\exp\Big[u\big(\pi_2(d=k/100,r=0),\pi_1(d=k/100,r=0)\mid x\big)\Big]\\ &+\sum_{k=1}^{100}\exp\Big[u\big(\pi_2(d=0,r=k/100),\pi_1(d=0,r=k/100)\mid x\big)\Big]\end{aligned}$$

in both cases.

Given decision data $y_1, y_2, \ldots, y_N$ from N responders as well as the corresponding proposer decisions $x_1, x_2, \ldots, x_N$, the log-likelihood function is

$$\log L = \sum_{i=1}^{N} \log \Pr(y_i \mid x_i).$$

We estimate the parameter b for each treatment while fixing the curvature parameter ρ at eight distinct values, so that the logarithm of the elasticity of substitution, $-\log(1-\rho)$, takes the values in $\{0.0, 0.5, 1.0, 1.5, 2.0, 3.0, 4.0, \infty\}$.

Table IV.4 presents the estimated values of b for these eight values of ρ across four treatments. The log-likelihood is maximised at $\rho = 0.0$; consequently, we focus our discussion on the results at this value.

The estimated values of b are positive only in the PTG treatment (although this estimate is not statistically significantly different from zero) and negative in the other three treatments. Moreover, in the PTG+CLR and PTG+G+CLR treatments, the estimates are statistically significantly different from zero.[4] These findings indicate that, when responders have the option to retaliate costlessly, they react more negatively—that is, they are more willing to reduce the proposer's payoff to increase their own—when proposers take more (or, in PTG+G+CLR, give less). This effect disappears when responders lack a costless-retaliation option.

[4]The estimated values of b are not significantly different between PTG and PTG+G ($p = 0.549$).

Table IV.4: Estimation results for the responders' parameter b

$-\log(1-\rho)$	ρ	PTG	PTG+CLR	PTG+G	PTG+G+CLR	$-\log L$
0.0	0.000	5.076	−9.226***	−2.560	−7.538**	
		(12.34)	(3.00)	(2.96)	(3.51)	
		176.68	*175.04*	*188.33*	*187.14*	*727.19*
0.5	0.393	20.511	−24.766***	1.040	−19.172**	
		(14.95)	(8.21)	(6.49)	(9.27)	
		180.74	*179.85*	*179.79*	*192.74*	*733.13*
1.0	0.632	11.838	−14.915**	0.854	−11.332	
		(12.67)	(6.45)	(4.12)	(6.78)	
		179.00	*180.81*	*178.40*	*192.72*	*730.93*
1.5	0.777	9.251	−11.830**	0.794	−8.875	
		(11.12)	(5.40)	(2.74)	(6.13)	
		178.30	*181.61*	*177.95*	*192.89*	*730.75*
2.0	0.865	8.115	−10.443*	0.761	−7.770	
		(11.04)	(5.30)	(2.63)	(6.44)	
		177.97	*182.15*	*177.77*	*193.04*	*730.93*
3.0	0.950	7.213	−9.326**	0.731	−6.881	
		(9.59)	(4.23)	(2.15)	(5.91)	
		177.70	*182.71*	*177.64*	*193.21*	*731.26*
4.0	0.982	6.922	−8.961*	0.719	−6.591	
		(9.22)	(5.25)	(2.12)	(5.52)	
		177.62	*182.92*	*177.60*	*193.27*	*731.42*
∞	1.000	6.761	−8.758*	0.713	−6.430	
		(10.58)	(4.51)	(1.96)	(5.91)	
		177.57	*183.05*	*177.58*	*193.32*	*731.52*

Note: Fixing the curvature parameter at eight distinct values, we estimated the parameter b for each treatment, where $\theta = b(t-g)$. In each cell, the top row reports the estimated value of b, the middle row its standard error and the bottom row the negative log-likelihood. Standard errors were calculated as the standard deviation of the estimates from the bootstrapped samples. The first column of the table reports the logarithm of the elasticity of substitution. When $\rho = 0$ (i.e., the Cobb–Douglas utility function), the total negative log-likelihood is minimised, indicating the best fit. ***, ** ,*: statistically significantly different from zero at 1%, 5%, and 10% significance level.

V Emotions

V.1 Initial emotion

Table V.1 shows the means (the standard deviations) of the emotions measured at the beginning of the experiment. There are no significant differences in any of these emotions across the four treatments.

V.2 Emotion after the instruction and after playing the game

Table V.2 shows the emotion elicited after the instructions and comprehension quiz (but before roles were revealed and the game was played) in Panel (a), as well as the changes from the initial emotions in Panel (b). There are significant differences between the changes from the emotions elicited at the beginning of the experiment and those after the instructions for the PTGs. In particular, participants in the PTG treatment exhibit significant increases in the intensity of irritation, anger, contempt, envy, jealousy, and sadness; no such increases occur in the other treatments, except for anger in PTG+G+CLR. Indeed, with respect to envy, the changes in the other three treatments are in the opposite direction (and are significant in PTG+CLR and PTG+G). The reasons for these differences across treatments remain unclear. Changes in joy, fear, and surprise are all in the same direction and significant across all four treatments, except for surprise in PTG+CLR.

Table V.1: Mean of initial emotions

	irritate	anger	contempt	envy	jealous	sad	joy	shame	fear	surprise
PTG	2.384	1.733	1.419	1.895	1.721	1.733	2.744	1.593	1.977	1.360
	(1.757)	(1.426)	(1.079)	(1.503)	(1.378)	(1.202)	(1.639)	(1.231)	(1.439)	(0.880)
PTG+CLR	1.915	1.463	1.341	1.939	1.707	1.744	2.695	1.634	1.939	1.451
	(1.363)	(1.021)	(0.892)	(1.443)	(1.356)	(1.294)	(1.600)	(1.171)	(1.534)	(0.772)
PTG+G	1.788	1.350	1.250	1.637	1.387	1.450	2.325	1.225	1.500	1.363
	(1.240)	(0.765)	(0.703)	(1.105)	(0.819)	(0.940)	(1.589)	(0.595)	(1.019)	(0.903)
PTG+G+CLR	2.013	1.387	1.337	1.525	1.475	1.613	2.388	1.450	1.738	1.462
	(1.514)	(0.961)	(0.941)	(0.968)	(1.018)	(1.268)	(1.626)	(0.953)	(1.329)	(1.158)
p-value (KW test)	0.368	0.710	0.829	0.649	0.821	0.544	0.177	0.225	0.265	0.567

Note: Standard deviations are shown in parentheses.

Table V.2: Mean of emotions after instruction and their changes from the initial emotion

(a) Emotions after the instruction

	irritate	anger	contempt	envy	jealous	sad	joy	shame	fear	surprise
PTG	2.698	2.267	1.767	2.209	1.884	2.442	2.279	1.465	3.023	2.640
	(1.873)	(1.798)	(1.485)	(1.790)	(1.648)	(1.766)	(1.547)	(1.165)	(2.023)	(1.928)
PTG+CLR	2.049	1.732	1.341	1.756	1.488	1.720	2.146	1.354	2.256	1.841
	(1.448)	(1.343)	(0.864)	(1.320)	(1.021)	(1.317)	(1.596)	(0.948)	(1.514)	(1.310)
PTG+G	2.013	1.512	1.325	1.462	1.312	1.538	1.900	1.200	1.975	1.788
	(1.268)	(1.019)	(0.742)	(0.980)	(0.908)	(1.179)	(1.308)	(0.537)	(1.350)	(1.366)
PTG+G+CLR	2.125	1.500	1.438	1.425	1.375	1.562	2.075	1.425	2.250	1.725
	(1.702)	(1.232)	(1.041)	(0.883)	(0.786)	(1.004)	(1.430)	(0.978)	(1.650)	(1.359)
p-value (KW test)	0.090	0.009	0.431	0.047	0.349	0.001	0.527	0.707	0.009	0.007

Note: Standard deviations are shown in parentheses.

(b) Changes from the initial emotion

	irritate	anger	contempt	envy	jealous	sad	joy	shame	fear	surprise
PTG	0.314**	0.535***	0.349**	0.314	0.163	0.709***	−0.465***	−0.128*	1.047***	1.279***
	(1.574)	(1.671)	(1.461)	(1.449)	(1.362)	(1.563)	(1.378)	(1.225)	(2.011)	(1.819)
PTG+CLR	0.134	0.268*	0.000	−0.183	−0.220	−0.024	−0.549***	−0.280***	0.317**	0.390**
	(1.245)	(1.415)	(0.969)	(1.145)	(1.122)	(1.176)	(1.188)	(0.725)	(1.498)	(1.377)
PTG+G	0.225	0.163**	0.075	−0.175**	−0.075*	0.087	−0.425***	−0.025	0.475***	0.425**
	(1.222)	(0.863)	(0.708)	(0.725)	(0.671)	(0.996)	(0.965)	(0.420)	(1.396)	(1.412)
PTG+G+CLR	0.113	0.113	0.100	−0.100**	−0.100	−0.050	−0.312***	−0.025	0.512***	0.263*
	(1.414)	(0.675)	(0.739)	(0.565)	(0.628)	(0.926)	(1.197)	(0.914)	(1.526)	(1.290)
p-value (KW test)	0.552	0.210	0.507	0.151	0.749	0.002	0.854	0.227	0.396	0.001

Note: ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively, using the Wilcoxon test. Standard deviations are shown in parentheses.

Table V.3 emotions from before to after playing the game for proposers (Panel (a)) and the responders (Panel (b)).[5]

As shown in Table V.3, we observed a significant difference in how emotions changed before and after the game between proposers and responders in the PTG and PTG+G treatments. For example, while the intensity of negative emotions—irritation, anger, and sadness—declined for the proposers after the game, it increased for the responders in these two treatments. Responders also reported significantly stronger feelings of contempt and jealousy after the game than before. In contrast, proposers did not experience reductions in these emotions in the remaining two treatments (PTG+CLR and PTG+G+CLR). In fact, in those treatments—where responders can engage in costless retaliation—we observed a significant increase in the intensity of contempt among proposers.

V.3 Principle component analysis of emotional changes

For each role (proposer and responder), several emotions exhibit similar directional changes across treatments. Therefore, we conduct separate principal component analyses for proposers and responders to summarise these changes.

The results are presented in Table V.4. The analyses reveal four prin-

[5]Table V.8 in Online Appendix V.2 reports only the emotions elicited after playing the game.

Table V.3: Mean of changes in emotions before and after playing the game

(a) Proposer

	irritate	anger	contempt	envy	jealous	sad	joy	shame	fear	surprise
PTG	−0.977***	−0.581***	−0.140**	−0.488***	−0.279***	−0.512**	2.326***	0.814***	−1.023***	0.674
	(1.406)	(1.180)	(1.060)	(0.960)	(0.797)	(1.387)	(2.032)	(1.500)	(2.076)	(2.254)
PTG+CLR	0.244	0.171	0.537**	−0.317***	−0.122*	0.293	0.634**	0.585***	−0.659***	0.707**
	(1.392)	(1.181)	(1.535)	(0.820)	(0.872)	(1.365)	(1.881)	(1.224)	(1.527)	(2.028)
PTG+G	−0.200	0.025	0.325*	0.025	0.150	0.100	1.725***	0.225	−0.500**	0.075
	(1.091)	(1.000)	(1.095)	(0.947)	(0.864)	(1.429)	(2.264)	(1.143)	(1.485)	(1.927)
PTG+G+CLR	−0.075	0.450*	0.350*	0.000	−0.025	0.150	0.800***	−0.175	−0.750***	0.625*
	(1.071)	(1.584)	(1.272)	(0.555)	(0.577)	(1.252)	(1.951)	(0.712)	(1.463)	(1.659)
p-value (KW test)	0.002	0.009	0.037	0.084	0.521	0.125	< 0.001	0.002	0.811	0.633

Note: ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively, using the Wilcoxon test. Standard deviations are shown in parentheses.

(b) Responder

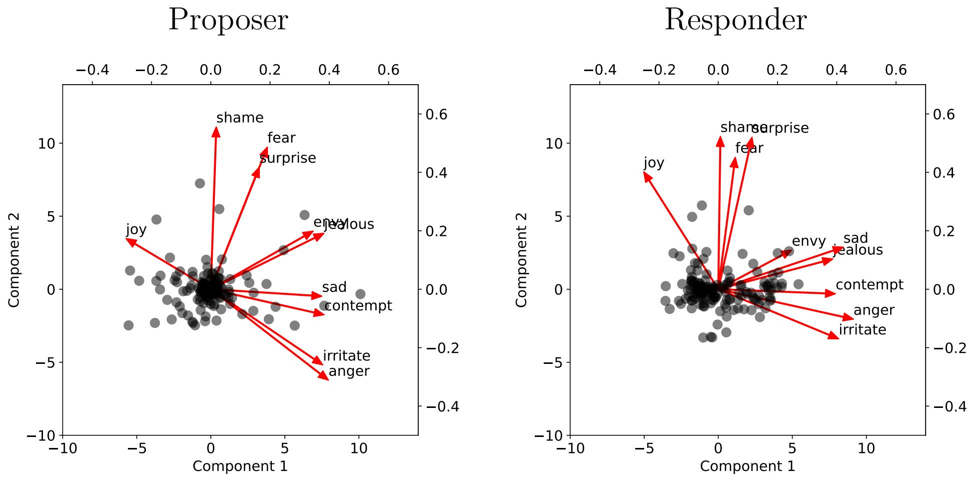
	irritate	anger	contempt	envy	jealous	sad	joy	shame	fear	surprise
PTG	0.977***	1.070***	1.116***	0.442*	0.977***	1.000***	0.070	0.023	−1.233***	−0.047
	(1.982)	(2.063)	(1.789)	(1.297)	(2.018)	(1.890)	(1.869)	(1.080)	(1.784)	(2.047)
PTG+CLR	1.171***	1.463***	1.390***	0.610**	0.878***	1.317***	0.000	−0.268	−0.707***	0.610**
	(2.011)	(2.259)	(2.060)	(1.412)	(1.503)	(2.138)	(1.688)	(1.119)	(1.504)	(1.896)
PTG+G	1.275***	1.875***	1.675***	0.925	1.675***	1.900***	−0.325**	0.150	−0.450**	0.450
	(2.276)	(2.323)	(2.464)	(2.291)	(2.291)	(2.827)	(0.944)	(0.975)	(1.616)	(2.342)
PTG+G+CLR	0.200	0.875**	1.175***	0.525	1.025***	0.925*	0.400	0.500	−0.450**	1.150**
	(2.186)	(2.090)	(2.217)	(1.679)	(1.993)	(2.212)	(2.580)	(1.826)	(1.632)	(2.445)
p-value (KW test)	0.038	0.169	0.775	0.949	0.432	0.246	0.355	0.738	0.346	0.165

Note: ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively, using the Wilcoxon test. Standard deviations are shown in parentheses.

(c) P-values for comparing proposer and responder (MW test, two-tailed)

	irritate	anger	contempt	envy	jealous	sad	joy	shame	fear	surprise
PTG	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.002	0.548	0.210
PTG+CLR	0.065	0.008	0.090	< 0.001	< 0.001	0.027	0.024	< 0.001	0.645	0.996
PTG+G	< 0.001	< 0.001	0.003	0.194	< 0.001	0.001	< 0.001	0.987	0.933	0.966
PTG+G+CLR	0.944	0.419	0.083	0.262	0.004	0.180	0.311	0.223	0.653	0.449

Figure V.1: Biplot of proposer's (left) and Player2's (right) emotional changes



Note: Arrows show the loadings (the top and right axes represent the scales), such that sum of squared values of the ten emotions equals one. Each dot represents a participant's score (the bottom and left axes represent the scales).

cipal components that together account for over 70% of the total variance. However, following Afifi et al. (2019, Ch. 14), we retain only the first two components, because the eigenvalues of the third and subsequent components are relatively small and level off compared to those of the first two components for both roles.

Figure V.1 shows the loadings of ten emotions on the two components for the proposer (left) and the responder (right), as indicated by arrows. We observe that, for both roles, the second component is driven primarily by changes in shame, fear, and surprise, whereas the first component is driven mainly by the remaining seven emotions, with joy loading in the opposite direction.

Table V.4: Summary of principal component analysis for emotion changes before and after playing the game

(a) Proposer

Variable	Component 1	Component 2	Component 3	Component 4
Δirritate$_{32}$	**0.378**	−0.259	−0.054	0.094
Δanger$_{32}$	**0.397**	**−0.311**	−0.182	0.206
Δcontempt$_{32}$	**0.383**	−0.088	−0.125	**0.481**
Δenvy$_{32}$	**0.346**	0.199	**0.553**	−0.168
Δjealous$_{32}$	**0.382**	0.192	**0.470**	−0.003
Δsad$_{32}$	**0.375**	−0.024	−0.002	−0.136
Δjoy$_{32}$	−0.286	0.174	**0.354**	**0.473**
Δshame$_{32}$	0.019	**0.556**	−0.032	**0.434**
Δfear$_{32}$	0.191	**0.487**	−0.259	**−0.479**
Δsurprise$_{32}$	0.164	**0.418**	**−0.478**	0.170
Eigenvalue	3.655	1.616	1.055	0.986
Cum. Prop.	0.366	0.527	0.633	0.731

(b) Responder

Variable	Component 1	Component 2	Component 3	Component 4
Δirritate$_{32}$	**0.407**	−0.170	−0.029	0.269
Δanger$_{32}$	**0.456**	−0.103	0.022	0.191
Δcontempt$_{32}$	**0.396**	−0.016	−0.006	0.246
Δenvy$_{32}$	0.248	0.134	**0.507**	**−0.521**
Δjealous$_{32}$	**0.387**	0.104	0.201	−0.274
Δsad$_{32}$	**0.422**	0.144	−0.068	−0.053
Δjoy$_{32}$	−0.252	**0.402**	**0.412**	0.203
Δshame$_{32}$	0.007	**0.523**	**0.333**	0.246
Δfear$_{32}$	0.058	**0.452**	**−0.554**	**−0.496**
Δsurprise$_{32}$	0.114	**0.521**	**−0.329**	**0.364**
Eigenvalue	3.779	1.555	1.082	0.906
Cum. Prop.	0.378	0.533	0.642	0.732

Note: Loadings with absolute values greater than 0.3 were highlighted.

Table V.5 reports the mean (standard deviation) of the two components for each role and treatment. For the proposer, the sign of Component 1 depends on whether the responder can engage in costless retaliation. In the treatments without this option (PTG and PTG+G), Component 1 is negative and significant in PTG. By contrast, in the treatments with costless retaliation (PTG+CLR and PTG+G+CLR), Component 1 is positive and significant in PTG+CLR. Component 2 for the proposer is positive and significant in PTG but negative in the other three treatments, reaching statistical significance only in PTG+G+CLR. Both components differ significantly across the four treatments ($p < 0.001$ for Component 1; $p = 0.015$ for Component 2, KW test).

For the responder, Component 1 is negative and significant only in PTG +G+CLR; it is positive—but not statistically significant—in the other three treatments.

Because we elicited emotions both before and after the game, these components capture the joint influence of proposers' and responders' . However, given the game's sequential structure, we can distinguish two directional effects. First, responders' emotional changes are driven primarily by proposers' decisions and, in turn, partly drive responders' own choices. Second, proposers' emotional changes are driven mainly by responders' actions. Indeed, Bosman and van Winden (2002) organise their analysis along very similar lines. In what follows, we adopt this sequential perspective in our data analysis.

V.3.1 Changes in emotions and s of responders

Table V.6 reports the results of regressing the two emotional components of the responder on the treatment dummies and the proposer's decision $(t-g)$. In Columns (2) and (4), we control for the responder's degrees of inequality aversion (α and β). In Column (1), we observe a positive and significant effect of the proposer's decision on Component 1 of the responder's emotions. Recall that Component 1 consists mainly of negative emotions, some of which are directed toward another person (for example, contempt, jealousy, and envy). Furthermore, there are no statistically significant differences across treatments.[6] These result remain unchanged even when we control for the degrees of inequality aversion (Column (3)). As Columns (5) and (6) show, however, Component 2 is not significantly related to the proposer's decision. This suggests that Component 2 captures changes in emotion that are not directly elicited by the proposer's action toward the responder.

We now analyse the emotions influencing the responder's decision. Table V.7 presents the result of two regressions: Column (1) uses the change in the proposer's payoff, and Column (2) uses the change in the responder's payoff as the dependent variable. Both regressions include the treatment dummies and the two emotional components as explanatory variables.

None of the estimated coefficients is statistically significant. The esti-

[6]None of the treatment dummies are significant. Additionally, none of the pairwise comparisons among PTG+CLR, PTG+G, and PTG+G+CLR are significant (Wald test).

Table V.5: Mean of two components

	Proposer		Responder	
	Component 1	Component 2	Component 1	Component 2
PTG	−0.999***	0.425*	−0.240*	−0.331**
	(1.658)	(1.615)	(1.687)	(1.252)
PTG+CLR	0.415*	−0.104	0.053	−0.173
	(1.917)	(1.244)	(1.782)	(0.904)
PTG+G	0.210	0.020	0.567	0.031
	(2.072)	(1.261)	(2.155)	(1.012)
PTG+G+CLR	0.438	−0.370***	−0.363	0.502
	(1.652)	(0.648)	(2.073)	(1.593)
p-value (KW test)	< 0.001	0.003	0.041	0.041

Note: ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively, using the Wilcoxon test. Standard deviations are shown in parentheses.

Table V.6: Effects on changes in emotions of responder in response to proposer's decision

	Component 1			Component 2		
	(1)	(2)	(3)	(4)	(5)	(6)
$I_{+\text{CLR}}$	0.29303	0.12167	0.09652	0.15820	0.17783	0.09845
	(0.42102)	(0.34480)	(0.35186)	(0.26594)	(0.26570)	(0.26990)
$I_{+\text{G}}$	0.80746*	0.59758*	0.62546*	0.36179	0.38584	0.31363
	(0.42370)	(0.34726)	(0.35443)	(0.26763)	(0.26760)	(0.27187)
$I_{+\text{G+CLR}}$	−0.12269	0.40269	0.39479	0.83231***	0.77212***	0.73558***
	(0.42370)	(0.35139)	(0.35739)	(0.26763)	(0.27078)	(0.27415)
$t - g$		0.02998***	0.03004***		−0.00344	−0.00314
		(0.00335)	(0.00337)		(0.00258)	(0.00259)
α			0.03816			−0.62289
			(1.30177)			(0.99855)
β			−0.27048**			−0.19415**
			(0.12811)			(0.09827)
const.	−0.24027	−1.84942***	−1.84941***	−0.33079*	−0.14642	−0.13233
	(0.29414)	(0.30016)	(0.30629)	(0.18579)	(0.23131)	(0.23495)
N	164	164	161	164	164	161
adj. R^2	0.01554	0.34175	0.35700	0.04565	0.05025	0.06455

Note: ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors are shown in parentheses.

Table V.7: Effects of changes in emotions of responder in response to proposer's decision

	Decrease in proposer's payoff due to responder's decision $(1-g)rY_1 + tdY_2$	Decrease in responder's payoff due to responder's decision $grY_1 + (1-t)dY_2$
	(1)	(2)
$I_{+\text{CLR}}$	511.128***	−4.488
	(67.031)	(7.915)
$I_{+\text{G}}$	69.997	14.120*
	(68.372)	(8.074)
$I_{+\text{G+CLR}}$	425.387***	5.061
	(70.089)	(8.277)
$t-g$	6.713***	−0.039
	(0.804)	(0.095)
Component 1	10.846	−0.373
	(15.450)	(1.824)
Component 2	15.899	−2.957
	(20.049)	(2.367)
const.	−264.742***	6.342
	(64.839)	(7.657)
N	164	164
adj. R^2	0.536	0.007

Note: ***, ** ,*: statistically significantly different from zero at 1%, 5%, and 10% significance level. Standard errors are shown in parentheses.

mated coefficients of the treatment dummies closely match those reported in Columns (2) and (4) in Panel (a) of Table 6.

Table V.8 reports the emotion elicited by players 1 (Panel (a)) and 2 (Panel (b)) after the game.

V.4 Emotional change of proposers

Finally, we examine the proposer's emotional response to the responder's decision. Table V.9 presents the results of regressing the first (Column (1)) and second (Column (3)) components of the proposer's emotional changes on treatment dummies and on the decreases in payoffs of players 1 and 2 resulting from the responder's decision. In Columns (3) and (6), we additionally control for the proposer's degree of inequality aversion (α and β).

Columns (2) and (3) show that, while the impact of the responder's action on the proposer's payoff ($(1-g)rY_1 + tdY_2$) positively and significantly influences the first component of the proposer's emotional changes, its impact on the responder's payoff ($grY_1 + (1-t)dY_2$) does not. Recall that the first component captures changes in negative emotions (including those directed toward the opponent). As hypothesised, the proposer's emotions are indeed influenced by the severity of the responder's reaction.

Moreover, the estimated coefficients of three treatment dummies—PTG+G, PTG+CLR, and PTG+G+CLR,—are all positive and significant. Thus, emotional reactions are significantly stronger in these treatments than in the

Table V.8: Mean of emotions after playing the game

(a) Proposer

	irritate	anger	contempt	envy	jealous	sad	joy	shame	fear	surprise
PTG	1.442	1.302	1.256	1.163	1.093	1.488	4.395	1.953	1.628	2.837
	(1.161)	(0.964)	(0.875)	(0.574)	(0.366)	(1.261)	(1.866)	(1.603)	(1.328)	(1.999)
PTG+CLR	2.366	1.976	1.902	1.463	1.366	1.976	2.927	1.829	1.610	2.561
	(1.813)	(1.620)	(1.578)	(1.098)	(1.090)	(1.557)	(2.138)	(1.321)	(1.321)	(1.747)
PTG+G	1.400	1.325	1.450	1.325	1.325	1.400	3.725	1.300	1.350	1.650
	(1.194)	(1.163)	(1.413)	(1.163)	(1.248)	(1.257)	(2.207)	(1.091)	(1.099)	(1.626)
PTG+G+CLR	1.625	1.800	1.525	1.375	1.250	1.425	2.750	1.250	1.300	2.225
	(1.275)	(1.713)	(1.339)	(0.952)	(0.840)	(1.083)	(1.836)	(0.840)	(0.758)	(1.804)
p-value (KW test)	0.020	0.058	0.349	0.530	0.861	0.101	0.001	0.010	0.757	0.005

Note: Standard deviations are shown in parentheses.

(b) Responder

	irritate	anger	contempt	envy	jealous	sad	joy	shame	fear	surprise
PTG	3.953	3.721	3.256	3.209	3.372	3.884	2.558	1.814	2.163	3.070
	(2.591)	(2.520)	(2.735)	(2.465)	(2.610)	(2.432)	(2.097)	(1.562)	(1.987)	(2.473)
PTG+CLR	3.146	3.122	2.707	2.341	2.366	3.073	2.000	1.195	1.537	2.439
	(2.545)	(2.619)	(2.390)	(1.970)	(2.059)	(2.392)	(1.466)	(0.511)	(1.306)	(1.718)
PTG+G	3.700	3.600	3.200	2.550	3.125	3.675	1.475	1.475	1.650	2.450
	(2.312)	(2.447)	(2.483)	(2.183)	(2.420)	(2.347)	(1.086)	(1.132)	(1.350)	(2.050)
PTG+G+CLR	2.750	2.525	2.875	2.000	2.500	2.775	2.600	1.925	2.000	3.000
	(2.216)	(2.219)	(2.388)	(1.783)	(2.124)	(2.315)	(2.085)	(1.900)	(1.840)	(2.298)
p-value (KW test)	0.076	0.041	0.665	0.243	0.170	0.085	0.083	0.469	0.496	0.644

Note: Standard deviations are shown in parentheses.

baseline treatment.[7] This result remains unchanged when we control for the degree of inequality aversion.

Component 2, by contrast, is not significantly related to respondents' decisions (Columns (5) and (6)). Recall that Component 2 consists mainly of surprise, shame, and fear. Component 2 scores are lower in PTG+G, PTG+CLR, and PTG+G+CLR than in the baseline treatment, and the differences are statistically significant for PTG+CLR and PTG+G+CLR.[8]

References

Afifi, A., May, S., Donatello, R. A. and Clark, V. A. (2019), *Practical Multivariate Analysis*, 6 edn, Chapman and Hall/CRC.

Bosman, R. and van Winden, F. (2002), 'Emotional hazard in a power-to-take experiment', *Economic Journal* **112**(476), 147–169.

Cox, J. C., Friedman, D. and Gjerstad, S. (2007), 'A tractable model of reciprocity and fairness', *Games and Economic Behavior* **59**, 17–45.

Csermely, T. and Rabas, A. (2016), 'How to reveal people's preferences:

[7]There are no significant differences across these three treatments. Based on Wald tests for Column (1), the p-values are 0.775 (PTG+G vs. PTG+CLR), 0.412 (PTG+G vs. PTG+G+CLR), and 0.604 (PTG+CLR vs. PTG+G+CLR).

[8]There are no significant differences across these three treatments. Based on Wald tests for Column (3), the p-values are 0.549 (PTG+G vs. PTG+CLR), 0.180 (PTG+G vs. PTG+G+CLR), and 0.480 (PTG+CLR vs. PTG+G+CLR).

Table V.9: Effects on changes in emotions of proposer in response to responder's decision

	Component 1			Component 2		
	(1)	(2)	(3)	(4)	(5)	(6)
$I_{+\text{CLR}}$	1.41387***	0.68509	0.68424	−0.52919*	−0.52934*	−0.55840*
	(0.39966)	(0.43277)	(0.44285)	(0.27271)	(0.30825)	(0.31658)
$I_{+\text{G}}$	1.20904***	1.03235***	1.05163***	−0.40501	−0.39032	−0.39658
	(0.40221)	(0.39271)	(0.39941)	(0.27445)	(0.27972)	(0.28553)
$I_{+\text{G+CLR}}$	1.43637***	1.01490**	0.98526**	−0.79472***	−0.78796***	−0.80379***
	(0.40221)	(0.40322)	(0.41517)	(0.27445)	(0.28721)	(0.29679)
Decrease in proposer's		0.00132***	0.00125***		−0.00001	0.00005
payoff, $(1-g)rY_1 + tdY_2$		(0.00035)	(0.00037)		(0.00025)	(0.00026)
Decrease in responder's		0.00030	−0.00050		−0.00107	−0.00068
payoff, $grY_1 + (1-t)dY_2$		(0.00389)	(0.00397)		(0.00277)	(0.00284)
α			1.14674			−0.21692
			(1.58013)			(1.12958)
β			−0.29503			0.14926
			(0.21403)			(0.15301)
const.	−0.99869***	−1.11567***	−1.07706***	0.42491**	0.43146**	0.42148**
	(0.27922)	(0.27133)	(0.27784)	(0.19053)	(0.19327)	(0.19862)
N	164	164	162	164	164	162
adj. R^2	0.08280	0.14844	0.14882	0.03421	0.02295	0.01650

Note: ***, **, and *: statistically significantly different from zero at 1%, 5%, and 10% levels, respectively. Standard errors are shown in parentheses.

Comparing time consistency and predictive power of multiple price list risk elicitation methods', *Journal of Risk and Uncertainty* **53**(2), 107–136.

Fehr, E. and Schmidt, K. M. (1999), 'A theory of fairness, competition, and cooperation', *The Quarterly Journal of Economics* **114**, 817–868.

Gonzalez, R. and Wu, G. (1999), 'On the Shape of the Probability Weighting Function', *Cognitive Psychology* **38**(1), 129–166.

He, H. and Wu, K. (2016), 'Choice set, relative income, and inequality aversion: An experimental investigation', *Journal of Economic Psychology* **54**, 177–193.

Inukai, K., Shimodaira, Y. and Shiozawa, K. (2024), Revisiting CES utility functions for distributional preferences: Do people face the equality–efficiency trade-off?, ISER Discussion Paper No. 1195R, Institute of Social and Economic Research, Osaka University.

Jacobson, S. and Petrie, R. (2009), 'Learning from mistakes: What do inconsistent choices over risk tell us?', *Journal of Risk and Uncertainty* **38**(2), 143–158.

Thöni, C. (2015), 'A note on CES functions', *Journal of Behavioral and Experimental Economics* **59**, 85–87.

Yang, Y., Onderstal, S. and Schram, A. (2016), 'Inequity aversion revisited', *Journal of Economic Psychology* **54**, 1–16.